

Sveučilište u Zagrebu  
Edukacijsko-rehabilitacijski fakultet

Diplomski rad

# Identifikacija govornika u forenzičkoj akustici i fonetici

Katarina Minarik

Zagreb, 2017

Sveučilište u Zagrebu  
Edukacijsko-rehabilitacijski fakultet

Diplomski rad

Identifikacija govornika u forenzičkoj akustici i fonetici

Katarina Minarik

Prof. dr. sc. Mladen Heđever

Zagreb, 2017

## Izjava o autorstvu rada

Potvrđujem da sam osobno napisala rad *Identifikacija govornika u forenzičkoj akustici i fonetici* i da sam njegova autorica.

Svi dijelovi rada, nalazi ili ideje koje su u radu citirane ili se temelje na drugim izvorima jasno su označeni kao takvi te su adekvatno navedeni u popisu literature.

Ime i prezime: Katarina Minarik

Mjesto i datum: Zagreb, 31.8.2017.

## Identifikacija govornika u forenzičkoj akustici i fonetici

Katarina Minarik

Prof. dr. sc. Mladen Heđever

Sveučilište u Zagrebu, Edukacijsko-rehabilitacijski fakultet, Odsjek za logopediju

### Sažetak

Forenzička akustika i fonetika relativno je novo interdisciplinarno područje koje objedinjavanjem tehničkih, biomedicinskih, humanističkih i društvenih znanosti utvrđuje relevantne činjenice u sudskim postupcima. Ovaj pregledni rad prikazuje jedno od područja rada kojim se forenzička akustika i fonetika bavi, a to je identifikacija govornika. Izum opreme za snimanje i reprodukciju zvuka, razni alati za akustičku obradu glasa te izum spektrograma, ključni su događaji koji su identifikaciju govornika učinili kakvom danas jest. Stručnjaci u području fonetike, lingvistike, patologije glasa i govora ili (elektro)akustike mogu obavljati glasovne usporedbe s ciljem identifikacije govornika, no sve to po određenim pravilima koje nalaže Međunarodno udruženje forenzičke fonetike i akustike (IAFPA). Dvije se metode procjene glasa i govora isprepliću i nadopunjuju, subjektivna/auditivna i objektivna/akustična. Posljednjih godina velika se važnost u ovom području daje pitanju: Je li naš glas uistinu jedinstven poput našeg otiska prsta? Pitanje jedinstvenosti glasa za sobom poteže i opise akustičkih obilježja glasa ( $F_0$ , jitter, shimmer, formanti) te obilježja govora (artikulacija, rezonancija, prozodija, obilježja jezika) koja naš glas i govor čine prepoznatljivima, a s druge strane i sve one čimbenike koji mogu utjecati na njegovu varijabilnost (stres, emocije, demografska obilježja, medicinsko stanje) uključujući i poremećaje logopedskog područja rada. U Hrvatskoj ne postoje stalni sudski vještaci u ovom području, niti specijalizirana ustanova pa tako ni propisane norme i procedure. Iako broj kaznenih djela za čije je razrješenje potrebno vještačenje glasa i govora raste već niz godina, kako u ostatku svijeta, tako i na području Hrvatske i regije, samo nekolicina osoba bavi se takvim vještačenjima u našoj zemlji. Stoga, cilj je ovog rada prikazati današnji status forenzičke akustike i fonetike u području vještačenja glasa i govora kroz pregled dosadašnjih znanstvenih i stručnih radova uz prikaz mogućnosti, ograničenja i novina koje ovo područje sa sobom nosi.

**Ključne riječi:** forenzička akustika i fonetika, identifikacija govornika, subjektivna i objektivna procjena glasa, glasovna varijabilnost.

## **Speaker identification in forensic acoustics and phonetics**

Katarina Minarik

Prof. dr. sc. Mladen Heđever

University of Zagreb, Faculty of Education and Rehabilitation Sciences, Department of  
Speech and Language Pathology

### **Summary**

Forensic acoustics and phonetics is a relatively new interdisciplinary area that determines relevant facts in court proceedings by encompassing technical, biomedical, humanities and social sciences. This review paper shows one of the areas that forensic acoustics and phonetics covers, and that is speaker identification. Invention of sound recording and reproduction equipment, various tools for acoustic voice manipulation and the spectrogram are key events that made the speaker identification what it is today. Experts in the areas of phonetics, linguistics, voice and speech pathology or (electro)acoustics can perform voice comparisons with the goal of speaker identification but everything under the set of rules that the International Association for Forensic Phonetics and Acoustics (IAFPA) dictates. Two methods of voice and speech assessment intertwine and complement each other, subjective/auditive and objective/acoustic. In recent years great attention in this area is given to the question: Is our voice truly unique like our fingerprint. The question of voice uniqueness ties to itself the descriptions of acoustic voice characteristics (F0, jitter, shimmer, formants) and speech characteristics (articulation, resonance, prosody, language characteristics) that make our voice and speech recognizable, and on the other hand all other characteristics that can influence its variability (stress, emotions, demographics, medical condition) also including speech pathology disorders. In Croatia there are no permanent court experts in this area, nor is there a specialized establishment, thus there are no set norms or procedures. Even though the number of felonies that could be solved by using expert testimony in this area is increasing, there are only a handful of individuals who perform these services in Croatia. Therefore the goal of this paper is to show the today's status of forensic acoustics and phonetics in the area of voice and speech expertise through an overview of current scientific papers, together with the possibilities, limitations and novelties in this area.

**Key words:** forensic acoustics and phonetics, speaker identification, subjective and objective voice evaluation, voice variability.

## SADRŽAJ:

1. UVOD .....	1
1.1. Određenje forenzičke akustike i fonetike i kratka povijest .....	1
1.2. IAFPA Kodeks struke vještaka forenzičkih fonetičara.....	6
1.3. Identifikacija govornika .....	7
1.3.1. Subjektivne i objektivne metode analize/obrade glasa i govora .....	10
1.3.1.1. Subjektivna/auditivna procjena .....	11
1.3.1.2. Objektivna/akustička procjena.....	12
1.3.1.2.1 MDVP (Multi Dimensional Voice Program) .....	13
1.3.1.2.2 PRAAT .....	18
1.4. Prednosti i ograničenja metoda za procjenu glasa.....	20
1.5. Akustika govornih zvukova .....	23
1.5.1. Fundamentalna frekvencija .....	23
1.5.2. Jitter i Shimmer .....	24
1.5.3. Formanti .....	25
1.6. Glotalni izvor glasa .....	28
1.7. Varijabilnost govornog signala .....	29
1.7.1. Emocije i govor .....	30
1.7.1.1. Sustavi za analiziranje stresa u glasu .....	31
2. CILJ RADA.....	33
3. OBILJEŽJA GOVORNIKA.....	34
1.1. Spol .....	34
1.2. Dob .....	35
1.3. Dijalekt.....	36
1.4. Strani naglasak.....	36
1.5. Sociolekt .....	37
1.6. Medicinsko stanje.....	37
4. ULOGA LOGOPEDA .....	38
5. NOVOSTI I BUDUĆA ISTRAŽIVANJA .....	39
6. FORENZIČKA AKUSTIKA I FONETIKA U HRVATSKOJ .....	41
7. ZAKLJUČAK.....	43

8. LITERATURA.....	44
Popis slika i tablica.....	1

# 1. UVOD

Ovaj pregledni diplomski rad obuhvaća samo dio onoga što postupak identifikacije govornika podrazumijeva. Za potrebe ovoga rada neće se ići u potankosti cjelokupnog postupka koji je poprilično kompleksan i koji, da bi bio kompletan, mora sadržavati i niz procedura koje prethode samoj glasovnoj usporedbi. Pod time se misli i na pripremu snimki, izradu kopija istih i poboljšanje snimki te ukoliko je potrebno i provjera autentičnosti, potom statistička obrada svih dobivenih podataka te naravno, uvjeti koje laboratorij za forenzičku akustiku i fonetiku treba ispunjavati. Nadalje, to uključuje i brojne zakonske regulative koje stoje u pozadini postupka identifikacije govornika, a i niz drugih pravila, osim onih spomenutih, koje Kodeks struke forenzičara-fonetičara nalaže. Također, navedeni logopedski poremećaji nisu jedini koje logopedska djelatnost obuhvaća, već su samo manji dio iste, a samim time niti jedini koji se mogu dovesti u vezu s forenzičkom akustikom i fonetikom u smislu vještačenja glasa, govora pa i jezika od strane stručnjaka logopeda. Ono što će ovaj rad obuhvatiti je kratko određenje forenzičke fonetike i ono najznačajnije što se dogodilo kroz povijest što je današnju forenzičku akustiku i fonetiku, a time i identifikaciju govornika doveli do razvoja i popularnosti. Spomenut će se jedno od najznačajnijih pitanja kada je riječ o postupku identifikacije, pitanje jedinstvenosti glasa, odnosno otiska glasa. Nadalje, stručnjaci kojeg profila te uz koje kvalifikacije i pod kojim pravilima mogu obavljati glasovne usporedbe. Zatim koje su metode upotrebljavanje u toj proceduri te koje su njihove prednosti, a koja ograničenja, uz detaljniji uvid u programe za akustičku obradu glasa, MDVP i PRAAT. Kako je govor dinamičan čin koji je pod utjecajem brojnih činitelja, kako vanjski, tako i unutarnjih, osobnih, objasnit će se koji su uzroci glasovne i govorne varijabilnosti i na koji način mogu mijenjati glasovne parametre i govorna obilježja. Neizostavni su i podatci o novinama u ovom području i smjernicama za buduća istraživanja te će na posljétku biti riječi o statusu forenzičke fonetike i akustike i vještačenju glasa i govora na području Hrvatske.

## 1.1. Određenje forenzičke akustike i fonetike i kratka povijest

Forenzička akustika i fonetika relativno je novo interdisciplinarno područje koje objedinjavanjem tehničkih, biomedicinskih, humanističkih i društvenih znanosti utvrđuje relevantne činjenice u sudskim postupcima. Područje rada forenzičke akustike i fonetike uključuje: a) utvrđivanje autentičnosti audio snimki (jesu li snimke naknadno "montirane", brisanje ili dodavanje sadržaja, dodavanje šumova ili zvukova kako bi se govor učinio



nerazumljivim), b) poboljšanje razumljivosti govora u audio snimkama, c) utvrđivanje okolnosti (mjesto i vremena) u kojima je snimka nastala (na temelju pozadinskih zvukova iz okoline), d) transkripcija razgovora, e) identifikacija (verifikacija) govornika, f) uzimanje iskaza kod osoba s teškoćama u komunikaciji (Heðever i Blaži, 2010). Sama akustika kao znanost vrlo je egzaktna, gledajući s aspekta fizike i matematike, no kada se govori o akustici govora, tada je ona jednako psihološke kao i fiziološke prirode. Akustička fonetika proučava akustičke osobine glasova i govora od kojih su glavne: amplituda, frekvencija, trajanje i spektar. S druge strane perceptivna ili auditivna fonetika proučava načine tumačenja glasova i govora. Pa se tako amplituda percipira kao glasnoća, frekvencija kao visina tona i trajanje kao duljina glasa (Heðever, 2009).

Identificirati osobu ili neki objekt značilo bi da je tu osobu ili objekt moguće razlikovati od svih drugih na Zemlji. Unutar forenzike ovaj proces bi se mogao nazvati individualizacijom te podrazumijeva redukciju od inicijalne populacije do pojedine osobe (Meuwly i Drygajlo, 2001). Glasovna identifikacija seže toliko daleko u povijest da toga nismo niti svjesni te se pretpostavlja da se već tisućama godina upotrebljava u pronalasku počinitelja nekog zločina kojega je netko auditivno, ali ne i vizualno percipirao. Prvi poznati slučaj pogrešne identifikacije govornika datira sve do 1660. kada je William Hulet bio osuđen za ubojstvo Kralja Charlesa I. pod svjedočanstvom Richarda Gittensa (Eriksson, 2005). Nekoliko ključnih izuma oblikovalo je forenzičku fonetiku kao što ona danas jest. Prvotno treba spomenuti izum telefona, zatim opreme za snimanje i reprodukciju zvuka, nakon toga razvoj raznih alata za akustičku analizu glasa te kao jednu od izrazito važnih stavki, spektrogram (Eriksson, 2005).

No, ako se govori o identifikaciji govornika učinjenoj od strane pravih stručnjaka, onda se može reći da je ta etapa započela s mogućnošću snimanja glasa na neki medij koji kasnije taj glas može reproducirati, što se veže uz razdoblje ranih 1930-ih. Također, uz sve veću upotrebu telefonskih linija povećava se i broj kaznenih djela učinjenih putem istih. Kako bi se takvi slučajevi pokušali razriješiti i kako bi postojala mogućnost analize tog snimljenog materijala, smatralo se da je taj akustički signal potrebno na neki način vizualizirati, čime dolazi do izuma spektrograma, što predstavlja prvi bitan korak k toj analizi. Ovdje valja spomenuti Bell Telephone Laboratories pri kojem je 1946. godine konstruiran uređaj pod nazivom Sonogram ili "visible speech analyzer" koji je printao akustičku energiju govornog signala prikazujući tako različite frekvencije glasa u određenom vremenskom trajanju (Lindh,

2017). Snimljena riječ ili rečenica mogla se beskonačno uzastopno ponavljati jer je uređaj imao beskonačnu magnetofonsku traku ili magnetski disk te se snimljena riječ i/ili rečenica ponavljala onoliko puta koliko je bilo filtara, a sam uređaj imao je cijelu banku pojasno-propusnih filtara (Heðever, 2009). Više o sonogramu će biti u kasnijim odlomcima.

Generalno gledajući u procesu identifikacije govornika sudjeluju i isprepliću se dvije grane te se mogu razlikovati fonetički i inženjerski pristup, a svaki od pristupa koristi se drugačijom metodologijom. Inženjerski pristup koji podrazumijeva automatske metode koristi znatno veći broj snimaka (glasa i govora) te podataka u svojoj analizi, a sve to uz više holističan pristup. S druge strane, fonetički pristup usmjeren je na mnoge ponašajne karakteristike individualnog govornika i daje na važnosti detaljima, no samim time koristi se mnogo manjom količinom podataka (Lindh, 2017).

Način na koji se računalno analiziraju govorni zvukovi čini se usko povezanim s načinom kako slušatelj ili, bolje rečeno, ljudsko uho analizira glas, čak i u novorođenačkom razdoblju ljudskog života (DeCasper i Spence, 1986). Dokazi upućuju na prepoznavanje, odnosno identificiranje (iako nesvjesno) majčina glasa pri samom rođenju djeteta, što upućuje da je takva identifikacija zapravo primarna funkcija naše slušne percepcije pri rođenju te da je ta funkcija tada bitnija od samog razumijevanja jezika i komunikacije (DeCasper i Fifer, 2004; prema Lindh, 2017).

Prema Kovač (2012, str. 349) „identifikacijom govornika mogu se baviti stručnjaci koji imaju primarno akademsko obrazovanje u području fonetike, lingvistike, patologije glasa i govora, odnosno logopedije ili (elektro)akustike te stručno iskustvo i znanje u forenzičkoj analizi glasa, govora i jezika.“ Forenzička identifikacija govornika dio je forenzičke fonetike te se najjednostavnije može objasniti kao postupak u kojem se kombinacijom auditivnih i akustičkih metoda pokušava prepoznati govornika u kaznenim postupcima (Rose, 2002). Ovaj je postupak po prvi puta implementiran 1980. godine u njemačkoj instituciji *Bundeskriminalamt* (Njemačka Federalna Policija) (Kunzel, 1995), a 1989. godine u Ujedinjenom Kraljevstvu održana je prva konferencija s temom primjene fonetike unutar forenzike.

Hollien (1990) u svojoj knjizi govori kako problematika vezana uz forenzičku fonetiku postaje sve raširenija te da bi uskoro to područje moglo biti od iste važnosti kao i druga područja unutar forenzičke znanosti.

## Otisak glasa ili *voice print*

Kako se ljudi često mogu susresti s ovom tematikom kroz televizijske emisije, filmove, kriminalističke romane i sl., dolazi do pojednostavnjivanja ovog podosta kompleksnog postupka forenzičke identifikacije govornika. Primjerice, ljudi mogu zaključiti da je identifikacija govornika zapravo nepogrešiva te u potpunosti pouzdana, da se u laboratorijskim uvjetima vrlo lako može identificirati govornika ili da je „otisak glasa“ (eng. *voice print*) ekvivalentan otisku prsta (Hollien, 2002). To bi značilo da je glas jedne osobe u potpunosti jedinstven i neponovljiv, a upravo je to dovelo forenzičku fonetiku do popularnosti, a posljedično i do njezina razvoja (Heðever, 2009). Kersta, (1962) dobiva rezultate koji govore u prilog gotove nepogrešivosti otiska glasa te navodi točnost u identifikaciji govornika od čak 98% putem metode vizualne usporedbe spektrograma glasa. Istraživanje Younga i Campbella (1967) iznosi podosta drugačije rezultate te oni govore da točnost identifikacije govornika vizualnom usporedbom otiska glasa u eksperimentalnim uvjetima iznosi svega 37,3% i to zbog znatnog utjecaja konteksta na govor pojedinca. Već tada počinje biti jasno da je identifikacija putem otiska glasa itekako sklona pogreškama. No, podijeljenost istraživača na one koji zagovaraju gotovu nepogrešivost identifikacije temeljem otiska glasa (Koenig, 1986) i onih koji takav stav kritiziraju (Ship i Doherty., 1987) nastavila se sve do kasnih 1980-ih godina. Vjeruje se da su zaposlenici Bellovog Laboratorija, koji je bio financiran od strane vojske, potajno razvijali metode prikazivanja „otiska glasa“. Tome je pridonijelo i to što njihovi istraživači u svojim radovima nisu spominjali ideju otiska glasa, već su svoj rad predstavljali kao fonetičko istraživanje koje bi trebalo uroditi plodom u području edukacije i terapije (Eriksson, 2005)

Metoda je kritizirana i od strane autora Bolt i sur. (1973) te oni u tom smislu postavljaju nekoliko pitanja:

- 1) Kada su dva spektrograma slična jedan drugom, znači li to da se radi o istom govorniku ili se radi o izgovoru iste riječi?
- 2) Mogu li irelevantne sličnosti odvesti sudsku porotu u krivom smjeru?
- 3) Koliko su uopće trajni/nepromjenjivi ti uzorci glasa?
- 4) Koliko su ta obilježja distinktivna ?
- 5) Mogu li se ta obilježja uspješno prikriti ili na neki način lažirati?

## Spektrogram/Sonogram

Zvukovni spektrogram uređaj je koji kao produkt daje vizualni prikaz govora, odnosno graf pod nazivom spektrograf koji prikazuje govor u funkciji vremena, frekvencije i glasovne energije (Koenig, 1986). Vrijeme trajanja govornog signala u sekundama je prikazan na x-osi, a govorne se frekvencije u hercima (Hz) nalaze na y-osi, dok je energetska komponenta govora izražena preko površine u nijansama od bijele do crne boje koja predstavlja energetski najizraženije dijelove govora (Kovač, 2012). Ta energetska najizraženija područja govora su frekvencije koje nazivamo formantima koji se pojavljuju pri izgovoru vokala (a, e, i, o, u) i nazala (m, n) te kod pojedinog govornika predstavljaju stalnu vrijednost, budući da odražavaju stanje vokalnog trakta i način artikulacije, odnosno obilježja prolaza zračne struje kroz vokalni trakt (Kovač, 2012), što ih čini važnima u postupku identifikacije govornika (prva tri formanta - F1, F2, F3). Najniža horizontalna linija je baza frekvencija i tu se obično nalazi trag osnovnog laringalnog tona (Heđever, 2009). Spektrogramska analiza sastavni je dio akustičkih programa za obradu glasa, no usporedba sonograma s ciljem identifikacije govornika svrstava se u subjektivnu metodu procjene govornika jer se oslanja na stručnjakovu vizualnu percepciju. Govoreći o hrvatskom jeziku, iskusan stručnjak vrlo jasno razlikuje različite govorne glasove prikazane spektrogramom, na primjer:

***Okluzivi*** - vidljiv je prazan bijeli prostor prije otpuštanja tlaka, a za vrijeme otpuštanja nastaje kratkotrajan intenzivan šum zvučne energije (Heđever, 2009).

***Frikativi*** - njihova se energija može uočiti kao aperiodičan signal (šum) koji nastaje raspršenjem energije na vrhu spektrograma, a trajanje i frekvencija frikativa su ključne značajke na spektrogramu. Raspršenje glasovne energije i frekvencijske karakteristike mogu ponekad biti ključni u prepoznavanju govornika koji imaju poremećaj izgovora kao što je sigmatizam (Heđever, 2009).

***Nazali*** - obilježja nazala na spektrogramu uključuju rezonantnu frekvenciju nosne šupljine te njen intenzitet. Ukoliko se nazalna komponenta uočava u spektrogramu i prilikom izgovora drugih ne-nazalnih glasova riječ je o poremećaju koji se naziva hipernazalnost te može biti važan faktor u prepoznavanju govornika (Heđever, 2009).

Rezonancija predstavlja temelj u prepoznavanju govornih uzoraka, a te su rezonantne karakteristike vidljive u spektralnom prikazu govora. Prema Heđever (2009; str.45)

„Promjene amplitude i frekvencije zvuka pomiješane s artikulacijskim promjenama nastalim u rezonatorima omogućavaju produkciju različitih glasova. Nekoliko složenih faktora ulaze u opservaciju generalnih rezonantnih karakteristika na spektrogramu: relativni intenzitet, disperzije energije i širina pojasa formanta te njihova razdvojenost“



Slika 1: Spektrogram (sonogram) (Preuzeto iz Heðever, 2012).

## 1.2. IAFPA Kodeks struke vještaka forenzičkih fonetičara

U sljedećim paragrafima izdvojene su neke od točaka Kodeksa struke vještaka forenzičkih fonetičara usvojenog na konferenciji Međunarodnog udruženja forenzičke fonetike i akustike 2004. godine u Helsinkiju (IAFP, 2004; preuzeto s <http://www.iafpa.net/code.htm>).

2. Identifikacijom raznolikog niza slučajeva koji su u domeni IAFP-a (to jest identifikacija/eliminacija govornika, profiliranje govornika, prepoznavanje glasova, transkripcije, svjedočanstva, pojačavanje audio signala, propagacija zvuka na mjestu zločina), članovi trebaju biti svjesni ograničenja vlastitog znanja i kompetencija kada pristaju na preuzimanje slučaja vještačenja.

5. Kod izvještaja o slučajevima u kojima se zahtijeva mišljenje ili zaključak, članovi trebaju jasno izraziti razinu sigurnosti i dati svoju procjenu u okviru skale identifikacije/eliminacije govornika.

6. (a) Članovi trebaju postupati s posebnim oprezom, ako izvode forenzičku analizu bilo kakve vrste na snimkama koje sadrže govor na jeziku koji nije njihov materinski jezik.

(b) Kod provođenja poslova forenzičke identifikacije/eliminacije govornika, članovi trebaju postupati s posebnim oprezom, ako su uzorci za uspoređivanje na različitim jezicima.

(c) Članovi trebaju postupati s posebnim oprezom, ako utvrđuju autentičnost ili ispituju integritet snimki za koje se sumnja da su originalni.

9. Članovi ne bi trebali izrađivati psihološke profile ili procjene iskrenosti govornika.

### **1.3. Identifikacija govornika**

Ljudi svakodnevno u raznim situacijama nesvjesno „analiziraju“ govornika. Svakoga puta kada se javimo na telefon ili čujemo u blizini nečiji glas i govor mi ga identificiramo te ukoliko se radi o osobi čiji glas dobro poznajemo dovoljna je sekunda da tu osobu prepoznamo temeljem samo njezina glasa i govora. Ponekad nam u tome pomažu i specifičnosti u govoru pojedinca, jezik kojim se služi, način na koji se izražava (Hollien, 2002) te brojna druga obilježja o kojima će biti riječi kasnije. No, u nekim situacijama identifikaciju nam otežava i sama okolina, odnosno prisutnost pozadinske buke koja se navodi kao najčešći problem u području forenzičke identifikacije govornika, a vezano je uz kvalitetu audio zapisa govornika (Eriksson, 2005). Zatim se navode odstupanja u govoru osobe do kojih može doći zbog snažnog utjecaja emocija ili prisutnosti bolesti. Nadalje, ako se radi o govorniku kojega čujemo putem telefonske linije ili putem raznih reproduksijskih uređaja, tada nam identifikaciju mogu otežati akustička obilježja koja su izgubljena u prijenosu signala (Hollien, 2002). U svakodnevnim situacijama navedeno ponekad i nije od velike važnosti jer ukoliko nismo sigurni tko je osoba s kojom razgovaramo imamo ju mogućnost to upitati. No, u slučajevima vještačenja glasa i govora sva su ta obilježja od značajne važnosti i nedostatak, odnosno gubitak ili maskiranje temeljnih akustičkih obilježja glasa i govora putem prijenosa signala može znatno otežati proces identifikacije govornika (Hollien, 2002). U slučaju kada izvorna snimka sadržava znatnu količinu buke ili ima ograničeni frekvencijski raspon, tada se kopija te snimke može pomoću različitih filtara unutar programa za obradu govornog signala pročistiti što za posljedicu ima bolju razumljivost govora i snimka je pripremljena za akustičku obradu (ponekad potrebna jedna kopija za preslušavanje, a druga za akustičku obradu) (Heðever, 2009).

Ni jedna osoba neće istu riječ izgovoriti dva puta s potpuno jednakim akustičkim i spektralnim obilježjima, tj. postoji nešto što znanstvenici koji se bave govorom nazivaju intra-govornikova varijabilnost. Suprotno tome, inter-govornikova varijabilnost odnosi se na

razlike u proizvodnji glasa/govora između različitih osoba te se temelji na razlikama u anatomskim strukturama koje sudjeluju u stvaranju glasa i naučenom načinu upotrebe govornog mehanizma (Tosi i sur., 1972). Sami govor neurološka je aktivnost koja započinje aktivnošću živčanog sustava, odnosno slanjem signala do različitih govornih organa zaduženih za produkciju. Nakon toga zrak iz pluća dolazi do larinksa gdje nastaju zvučni valovi zahvaljujući vibracijama glasnica. Taj se zvučni val nadalje prenosi kroz farinks u usnu ili nosnu šupljinu (/m/, /n/, /nj/) gdje poprima svoj konačan oblik prije no što izađe kao određeni glas preko usana ili nosa (Kvistedal, 2000). Oralna i nosna šupljina rezonatori su čiji oblik, veličina i stanje utječu na inter-individualna obilježja u proizvodnji glasova. Poznavanje strukture vokalnog trakta te razumijevanje njegove funkcije može biti presudno u rješavanju slučajeva u području forenzičke fonetike (Hollien, 1990). Stručnjaci koji uspoređuju snimke glasova govornika moraju biti svjesni utjecaja razlika u vokalnog trakta koje dovode do specifičnih razlika među govornicima. Također, ti se stručnjaci i dalje nadaju da će buduća istraživanja uzeti u obzir ovaj problem i doći do odgovora koji je najbolji način identifikacije govornika, a da se ova obilježja ne ispuste i zanemare (Bijhold i sur., 2007).

Svaka je forenzička analiza glasa i govora multidimenzionalna te obuhvaća saznanja o (Hedevar i Blaži, 2010):

1. respiraciji
2. fonaciji – glasu
3. artikulaciji - govoru
4. rezonanciji – vokalni trakt
5. anomalijama vokalnog trakta i bolestima
6. prozodiji
7. jeziku – dijalektu – poštapalicama
8. emocionalnom stanju i intoksikaciji tijekom govora

Sve se više počelo prihvaćati identificiranje metodom auditivne percepcije te se često ona izjednačavala s, odnosno shvaćala jednako pouzdanom metodom kao i vizualna identifikacija (od strane osobe koja je svjedočila zločinu). No, ono što tada nisu uzeli u obzir je činjenica da je pamćenje akustičkog signala sklono varijacijama te se po svojim karakteristikama razlikuje

od vizualnog pamćenja. (Hollien, 2002). Eriksson (2005) spominje kako je uobičajeno da prođe određeni vremenski period između počinjenog zločina te pokušaja donošenja odluke o tome odgovara li glas osumnjičenika samom glasu počinitelja. Upravo je iz toga razloga slušno pamćenje glasa i govora počinitelja od strane svjedoka kritično za ovaj dio postupka. Još 1937. godine Francis McGehee je u svom radu iznio podatke koji idu u prilog opadanju postotka točnosti identifikacije glasa nakon određenog vremenskog razmaka od prvotnog slušanja govornika. Ispitanici su slušali govornika koji je čitao tekst od 56 riječi, nakon čega su bili raspodijeljeni u grupe koje su u različitim vremenskim razmacima opetovano čuli istog govornika te ga trebali prepoznati u takozvanom poretku glasova (eng. *voice line-up*). Nakon razmaka od jednog dana do tjedan dana postotak točne identifikacije iznosio je 80%. Nakon samo dva tjedna postotak pada na 69%. Razmak od mjesec dana dovodi do točne identifikacije u 57% slučajeva, nakon tri mjesec iznosi svega 35%, a u slučajevima kada od prvog slušanja do pokušaja identifikacije prođe pet mjeseci, vjerojatnost točne identifikacije iznosi svega 13%. Na točnost identifikacije utjecat će i iskustvo onoga koji sluša, sličnost glasova različitih govornika na audio snimkama, duljina snimke/uzorka govora i naravno kvaliteta snimke (Eriksson, 2005).

Postoje određeni kriteriji koje ispitivač treba ispunjavati da bi se smatrao pouzdanim za obavljanje glasovne usporedbe s ciljem identifikacije. FBI (*Federal Bureau of Investigation*) u studiji iz 1985 (prema Koenig, 1986) navodi sljedeće kvalifikacije kod svojih zaposlenika ispitivača: minimum od 2 godine iskustva u radu na identifikaciji glasa koristeći sofisticiranu opremu, učinjeno preko stotinu usporedbi glasa u sklopu pravih forenzičkih slučajeva, sudjelovanje u dvotjednoj obuci za ovladavanjem spektrogramске analize, prolazak na godišnjem testu sluha, zatim je potrebno imati formalno dopuštenje za rad od strane drugih obučavanih ispitivača te posjedovati barem prvostupničku diplomu u području znanosti.

Vrlo je važno istaknuti da kada stručnjaci forenzike donose odluku o identitetu govornika ne mogu govoriti ni o kakvoj kategoričkoj odluci u smislu davanja izjave kako glas s neke konkretne snimke odgovara zasigurno konkretnoj osobi. Zbog toga postoji omjer vjerojatnosti (eng. *likelihood ratio*, LR) (Neyman i Pearson, 1933) kojim ti stručnjaci u svojim vještačenjima izražavaju sigurnost u svoju odluku (Morrison, 2009) koja nikada nije sto postotna identifikacija ili eliminacija, nego se radi o vjerojatnosti da glas određenog govornika pripada osobi sa snimke koja se analizirala (Hansen i Hasan, 2015). Kako bi ipak stručnjak mogao biti sigurniji u svoju odluku, a s obzirom na varijabilnost tijekom



dinamičnog govora, potrebno je uzeti što više uzoraka istog iskaza jer tada je i više valjan i pouzdan proces identifikacije glasa (Heđever, 2009).

### ***1.3.1. Subjektivne i objektivne metode analize/obrade glasa i govora***

Kako bi se izbjegao utjecaj okolne buke na snimci govornika prvo se odvija normalizacija signala koristeći se keprstralnom supstrakcijom, a time se mogu i reducirati razlike među signalima snimanim u različitim uvjetima (Lindh, 2017).

Procjena glasa/govora nepoznatog govornika na temelju slušanja i opažanja od strane stručnjaka, a potom donošenje odluke na temelju ljudskih misli naziva se subjektivnom analizom glasa. Za razliku od toga, analiza koja se vrši putem računalnih programa koji automatski daju podatke o akustičkim obilježjima glasa sa snimke poznata je kao objektivna analiza. No, i ovakva objektivna analiza zahtijeva krajnju interpretaciju vrsnog stručnjaka pri donošenju konačne odluke (Heđever, 2009).

U razrješavanju forenzičkih slučajeva, kako bi se identificirao nepoznati govornik upotrebljava se takozvano automatsko identificiranje govornika. No, novija literatura nalaže kako se više ne koristi u tolikoj mjeri termin „identifikacija“, već zbog probabilističke prirode ovakvih forenzičkih procedura ispravno bi bilo koristiti termin automatsko glasovno uspoređivanje (eng. *automatic voice comparison, AVC*). Takav sustav uspoređuje akustičke vektore glasova te koristi statističke mjere, ignorirajući lingvističke i/ili fonetičke informacije. Drugim riječima, to je biometrički izračun signala koji se istražuje. Budući da je ljudski glas podložan varijacijama u različitim situacijama, ovakva automatska analiza treba biti popraćena modelima koji će u snimci tog govornika obratiti pozornost i na fonetska, fonološka i lingvistička obilježja (Lindh, 2017).

Općenito, analiza se govornikova glasa i govora može podijeliti u 3 stadija: 1) uređivanje snimke te lingvističko fonetička analiza iste (ovdje spada odabir dijelova snimke koji će se upotrijebiti, transkribiranje te analiza lingvističkog ponašanja govornika sa snimke), 2) dobivanje akustičkih parametara kao što su fundamentalna frekvencija, govorna brzina te formanti i na posljetku 3) takozvana biometrijska usporedba glasovne kvalitete pomoću automatskih sustava, kod koje je bitno iskustvo onoga koji analizira. U ovoj etapi analize akustička se obilježja glasa osumnjičenika uspoređuju s onima sa postojeće snimke te se

dobiva rezultat koji ukazuje na sličnost tih dvaju glasova, odnosno vjerojatnost da se radi o istoj osobi (Lindh, 2017).

Akustička obilježja mogu se razdijeliti na kratkotrajna i na ona koja su dugotrajna. Automatski sustavi obrade/analize govornog signala uglavnom koriste kratkotrajna akustička obilježja. Na primjer, kako bi na direktan način usporedili glas /r/ između više govornika. Kada bi ta kratkotrajna obilježja uprosječili, tada bi dobili takozvana dugotrajna obilježja koja su manje osjetljiva na varijacije govornog signala neke individue. U tu skupinu spadaju na primjer, energija te govorna visina, koji su također sastavni dijelovi današnjih automatskih sustava analize (Hansen i Hasan, 2015).

#### **1.3.1.1. Subjektivna/auditivna procjena**

Subjektivna se metoda može podijeliti na dvije od kojih je jedna „više subjektivna“ od druge. Prva je već navedena, u kojoj iskusni slušatelj uspoređuje snimke poznatih i nepoznatih govornika pritom koristeći svoje kratkoročno i dugoročno pamćenje, a druga bi bila vizualna usporedba putem glasovnom spektrograma. Unutar objektivnih metoda se isto tako razlikuju dvije, poluautomatska i automatska (Tosi, 1979; prema Heđever, 2009).

Prema Tosi (1979; prema Heđever, 2009) svaka subjektivna odluka može rezultirati jednim od sljedeća četiri (4) odgovora:

1. Vrlo nesiguran da je moja odluka točna
2. Prilično nesiguran da je moja odluka točna
3. Prilično siguran da je moja odluka točna
4. Skoro siguran da je moja odluka točna.

U sljedećoj tablici navedena su obilježja koja su auditivno uočljiva te koja se opisuju pri subjektivnoj procjeni glasa i govora.

**Tablica 1: Auditivno uočljiva obilježja glasa i govora uz pripadajuće kategorije**

	<b>Obilježje</b>	<b>Kategorije</b>
<b>1.</b>	VISINA GLASA	1 –niski; 2- srednji; 3-visoki
<b>2.</b>	PREKIDI GLASA	1- nema prekida; 2- povremeni; 3- učestali
<b>3.</b>	INTONACIJA	1- monotona 2- normalna; 3- pjevajuća
<b>4.</b>	ATAKA GLASA	1- normalan početak; 2 – šumna; 3 - tvrda
<b>5.</b>	DISFONIJA	1- nema; 2- blaga; 3-umjerena; 4- jaka
<b>6.</b>	PROMUKLOST	1- nema; 2- blaga; 3-umjerena; 4- jaka
<b>7.</b>	HRAPAVOST	1- nema; 2- blaga; 3-umjerena; 4- jaka
<b>8.</b>	ŠUMNOST	1- nema; 2- blaga; 3-umjerena; 4- jaka
<b>9.</b>	SLABOST GLASA	1- nema; 2- blaga; 3-umjerena; 4- jaka
<b>10.</b>	NAZALNOST (REZONANCIJA)	1- hiponazalan; 2- normalna rezonancija; 3 hipernazalan
<b>11.</b>	GLASNOĆA	1- tiši; 2- normalan; 3- glasniji govor
<b>12.</b>	BRZINA ARTIKULACIJE	1- sporija; 2- normalna; 3- "guta" glasove
<b>13.</b>	BRZINA GOVORA/TEMPO	1- sporija; 2- normalna; 3-ubrzana
<b>14.</b>	GOVORNI RITAM	1- pravilan ujednačen; 2- promjenjiv varijabilan
<b>15.</b>	DIJALEKT	1- standardan govor; 2- blagi; 3- jaki dijalekt

### **1.3.1.2. Objektivna/akustička procjena**

Svaka analiza može rezultirati samo jednom od sedam (7) odluka (Heðever, 2009):

1. Sigurna identifikacija,
2. gotovo sigurna identifikacija,
3. moguća identifikacija,
4. nedefinirana identifikacija (nema zaključka),
5. moguća eliminacija,

6. gotovo sigurna eliminacija i

7. eliminacija.

#### ***1.3.1.2.1 MDVP (Multi Dimensional Voice Program)***

Tvrtka KayPENTAX (kasnije nazvana Kay Elemetrics) stavlja na tržište 1993. godine računalni program za akustičku obradu glasa pod nazivom Multi Dimensional Voice Program (MDVP) koji je opcionalno sastavni dio dvaju programa iste tvrtke: CSL (Computer Speech Lab) i Multi-Speech (Nicastri i sur., 2004). Zahvaljujući tom programu od tada je bilo moguće analizirati 33 kvantitativna akustička parametra glasa, uključujući i fundamentalnu frekvenciju (F0), intenzitetske vrijednosti, spektralna obilježja te bilježiti promjene sonornosti i prisutnost diplofonije. Njime se dobiva robustna multidimenzionalna analiza glasa, a dobivene vrijednosti parametara prikazane su numerički i grafički. Dvije su verzije ovoga programa, osnovna i napredna, koje stručnjak upotrebljava ovisno o vlastitoj razini vještina. Sama kvaliteta ulaznog i izlaznog signala ovisi o kvaliteti same zvučne kartice računala koje se upotrebljava, a program se može pokrenuti na računalima s Microsoft Windows XP te Windows Vista operacijskim sustavom (Kay PENTAX, 2008).

#### **Obrada govornog signala**

Signal koji se želi analizirati može se otvoriti u programu s memorije diska ili ga snimiti u tom trenutku putem MDVP programa, a nakon što je obrada gotova program generira šest prozora svaki od kojih pokazuje rezultate različitih analiza, uključujući i grafički prikaz parametara. U jednom od šest prozora prikazane su vrijednosti fundamentalne frekvencije (F0) te Peak-to-Peak amplitude koje su dobivene izračunom iz glotalnog perioda za svaki ciklus titranja glasnica (KayPENTAX, 2008).

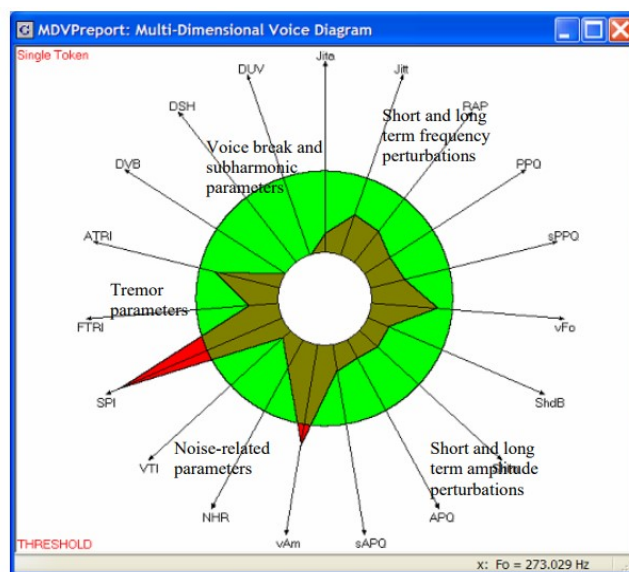


**Slika 2: Grafički prikaz fundamentalne frekvencije (F0) (crveni trag) i amplitude (plavi trag) između dva ciklusa, prikazano programom MDVP (Preuzeto iz KayPENTAX, 2008).**

Također, u ovom primjeru vidljivo je kako nije bilo velikih varijacija fundamentalne frekvencije, dok se isto ne može reći za intenzitet te se tu mogu uočiti amplitudne varijacije. Jednim od algoritama MDVP programa dobivaju se i spektralna obilježja signala putem srednjeg dugotrajnog spektra (eng. *Long Term Average Spectrum*, LTAS). Vrijednosti dobivene spektralnom analizom: indeks turbulencije glasa (VTI), omjer signal-šum (NHR) te indeks meke fonacije (SPI) prikazuju se u trećem prozoru programa. Zatim se u četvrtom i petom prozoru dobije prikazan histogram za vrijednosti F0 i amplitude, odnosno raspršenje vrijednosti tih parametara (Kay PENTAX, 2008).

Radijalni dijagram, koji se pojavljuje u šestom prozoru, numeričke vrijednosti MDVP analize signala prikazuje kružnim grafom. Svaki od parametara prikazan je linijom s početnom točkom u središtu kružnog dijagrama, a što je linija više udaljena od središnje početne točke ukazuje da se radi o povećanju vrijednosti tog parametra. Pragovi normativnih vrijednosti parametara vidljivi su putem zeleno obojenog kružnog područja grafa. Zatim su vrijednosti dobivene obradom vokalizacije prikazane tamnijom zelenom bojom ako se radi o vrijednostima unutar normativnih pragova ili crvenom bojom ukoliko su te vrijednosti izvan normativnih. Moguće je i dobiti prosječne vrijednosti izračunate iz nekoliko pokušaja vokalizacije. Ukoliko su dobivene vrijednosti izvan normi to ne znači nužno da je glasovna patologija zasigurno prisutna. MDVP je programiran tako da u slučaju da pokazuje da je neka dobivena vrijednost izvan normi, to radi sa sigurnošću od 95% (KayPENTAX, 2008).

Vrijednosti su na kružnom dijagramu organizirani po tipu parametra:



Slika 3: Radijalni dijagram MDVP programa s prikazom parametara (Preuzeto iz KayPENTAX, 2008).

Uobičajeno je da akustička obrada obuhvaća sljedeće korake: 1) profiliranje raspona glasa (pomoću fonetograma) – ovo se može učiniti pomoću CSL sustava putem VRP (*Voice Range Profiling*) programa, 2) analiza raspona visine i 3) analiza habitualne visine glasa (oboje putem *Real-Time Pitch* programa), zatim 4) spektrogramska analiza te 5) analiziranje glasovnih parametara (MDVP to čini na temelju produljene fonacije) (Kay PENTAX, 2008).

Kada se sve to učini, MDVP ima mogućnost napraviti izvještaj u kojem će stajati numeričke vrijednosti, grafički prikaz, informacije o pacijentu/osobi sa snimke (ako su poznate) te komentare ispitivača (KayPENTAX, 2008).

**MDVPreport: Voice Report**

Institution: St. Joseph's Hospital Date: Sep 23, 2005, Fri Acc.# 7235-2358

Name: Mary Brown Gender: Female Age: 45 File:

Address: 2 Bridgeport Avenue City: Geodora State: NJ ZIP: 07077

Diagnosis:

Comments:

Parameter	Name	Value	Unit	Norm(f)	STD(f)	Threshold
Average Fundamental Frequency	Fo	229.956	Hz	243.973	27.457	
Mean Fundamental Frequency	MFo	229.932	Hz	241.080	25.107	
Average Pitch Period	To	4.349	ms	4.148	0.432	
Highest Fundamental Frequency	Fhi	239.195	Hz	252.724	26.570	
Lowest Fundamental Frequency	Flo	224.477	Hz	234.861	28.968	
Standard Deviation of Fo	STD	2.331	Hz	2.722	2.115	
Phonatory Fo-Range in semi-tones	PFR	2		2.250	1.060	
Fo-Tremor Frequency	Fftr	3.846	Hz	3.078	1.964	
Length of Analyzed Sample	Tsam	3.750	s	3.000	0.000	
Absolute Jitter	Jita	19.530	µs	26.927	16.654	83.200
Jitter Percent	Jitt	0.449	%	0.633	0.351	1.040
Relative Average Perturbation	RAP	0.261	%	0.378	0.214	0.680
Pitch Perturbation Quotient	PPQ	0.267	%	0.366	0.205	0.840
Smoothed Pitch Perturbation Quotient	sPPQ	0.509	%	0.532	0.220	1.020

Save As... Print... Info... OK Cancel

Slika 4: MDVP izvješće dobivenih glasovnih parametara (Preuzeto iz KayPENTAX, 2008).

Tablica 2: Popis parametara unutar MDVP programa (Preuzeto iz KayPENTAX, 2008).

Oznaka i mjerna jedinica	Parametar
Fo, Hz	Average Fundamental Frequency
To, ms	Average Pitch Period
Fhi, Hz	Highest Fundamental Frequency
Flo, Hz	Lowest Fundamental Frequency
STD, Hz	Standard Deviation of Fo
PFR,	Phonatory Fo-Range in semi-tones
Fftr, Hz	Fo-Tremor Frequency
Fatr, Hz	Amplitude Tremor Frequency
Tsam, s	Length of Analyzed Sample
Jita, us	Absolute Jitter
Jitt, %	Jitter Percent

RAP, %	Relative Average Perturbation
PPQ, %	Pitch Perturbation Quotient
sPPQ, %	Smoothed Pitch Perturbation Quotient
vFo, %	Fundamental Frequency Variation
ShdB, dB	Shimmer in Db
Shim, %	Shimmer Percent
APQ, %	Amplitude Perturbation Quotient
sAPQ, %	Smoothed Amplitude Perturbation Quotient
vAm, %	Peak-to-Peak Amplitude Variation
NHR,	Noise to Harmonic Ratio
VTI,	Voice Turbulence Indeks
SPI,	Soft Phonation Indeks
FTRI, %	Fo-Tremor Intensity Indeks
ATRI, %	Amplitude Tremor Intensity Indeks
DVB, %	Degree of Voice Breaks
DSH, %	Degree of Subharmonics
DUV, %	Degree of Voiceless
NVB,	Number of Voice Breaks
NSH,	Number of Subharmonic Segments
NUV,	Number of Unvoiced Segments
SEG,	Number of Segments Computed
PER,	Total Number Detected Pitch Periods



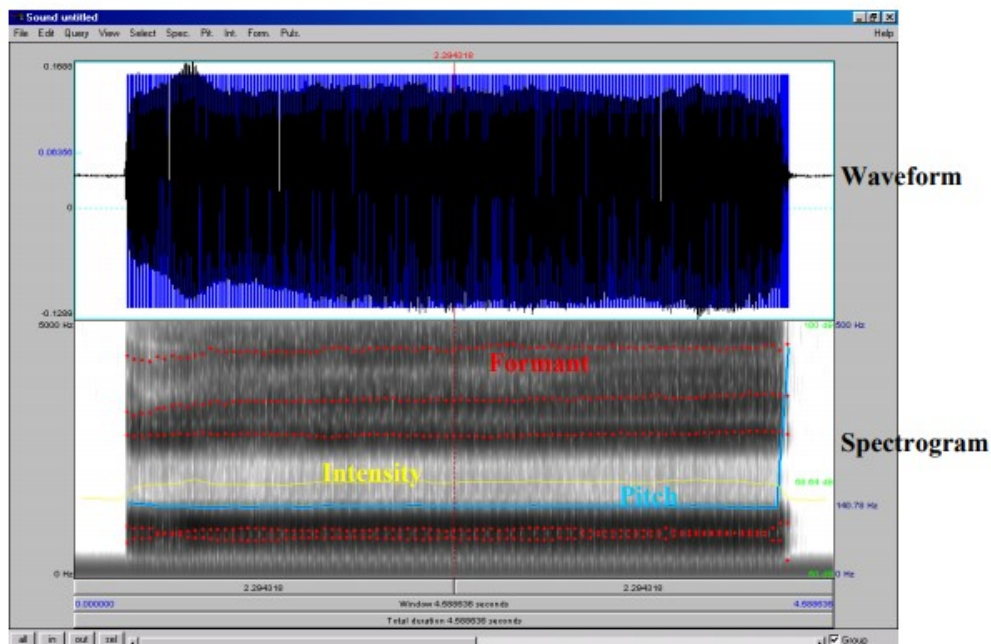
### **1.3.1.2.2 PRAAT**

PRAAT računalni program razvijaju Paul Boersma i David Weenink 1992. (Lieshout, 2003) godine na Institutu fonetskih znanosti pri Sveučilištu u Amsterdamu. Putem istoga moguće je analizirati govorni signal, njime manipulirati te stvarati sintetizirani zvuk. Moguće ga je potpuno besplatno preuzeti sa internetske stranice <http://www.fon.hum.uva.nl/praat/>. Ovaj je program vrlo fleksibilan alat za obradu glasa i govora te nudi široki raspon standardiziranih i ne standardiziranih procedura, uključujući spektrogramsku analizu, artikulaturnu sintezu te neuralne mreže (Lieshout, 2003).

Bilo da je neki zvuk snimljen putem mikrofona ili se zvuk prethodno snimljen otvara u programu PRAAT, ono daje mogućnost vidjeti unutar tog zvuka, odnosno njegovu reprezentaciju u obliku zvučnog vala (gornji dio programskog prozora), a isto tako i njegova akustička obilježja uključujući spektralni prikaz (donji dio programskog prozora). Kako se PRAAT najčešće koristi za analiziranje ljudskih glasova govora, tako nam pokazuje konture i vrijednosti visine glasa, fundamentalne frekvencije (brzina titranja glasnica) te vrijednosti i konture formanta (rezonancija vokalnog trakta), koji su ujedno osnovno obilježje spektrograma. Oba se prikaza (oscilogram i spektrogram) mogu uvećati i umanjiti, ovisno o obilježjima i detaljima koje želimo prikazati za određeni zvučni ili govorni signal. Istim se tim zvukovima može manipulirati, na primjer mogu se rezati dijelovi koji nisu relevantni, ne sadrže bitan sadržaj ili se radi o dijelovima bez govora pa taj dio želimo ukloniti iz snimke (Lieshout, 2003).

Inicijalno se pri pokretanju programa otvaraju dva programska prozora, lijevi pod nazivom „Praat objects“ te desni u kojem se prilikom obrade signala izbacuju grafovi (Kovač, 2012). Želimo li se baviti specifičnim frekvencijskim pojasom zvučnog signala, npr. želimo li pojedine frekvencije istaknuti, zvučne je objekte moguće obrađivati filtriranjem. Pod generalnu obradu zvučnog signala u PRAAT-u spadaju oscilogram, intenzitet, spektrogram, visina te trajanje, a što je također kompatibilno s prethodno spomenutim CSL softverom. Zatim je tu i opcija obrade glasa koja se vrši uz snimku produljene fonacije vokala /a/ u trajanju od najmanje šest sekundi. Ovdje se obradom dobivaju i vrijednosti frekvencijskih (jitter) te amplitudnih (shimmer) iregularnih varijacija. PRAAT program također ima mogućnost generiranja glasovnog izvještaja iz kojega se mogu iščitati vrijednosti svih izračunatih parametara. Još treba spomenuti da se kod spektrogramske analize, ovisno o tome želimo li jasniji prikaz formanta ili pojedinih harmonika, može odabrati između uskopojasne

(eng. *narrowband*) i širokopojasne (eng. *broadband*) analize. Ukoliko se odabere duljina trajanja prozora (eng. *window lenght*) u trajanju od 0.05 sekundi, što odgovara trajanju jednog perioda zvuka od 20 Hz, to znači da će u takvom spektrogramskom prikazu biti komponente višekratnika frekvencije od 20 Hz, a takav prikaz će tada biti primjeren za uočavanje harmonika. S druge strane, širokopojasni prikaz bolji je za uočavanje formanta te se u tom slučaju može izabrati duljinu trajanja prozora od 0.005 sekundi, a tada će u prikazu spektrograma biti komponente višekratnika frekvencije od 200 Hzkompletan.



**Slika 5: Oscilogram (gornji dio prozora) i spektrogram (donji dio prozora) dobiveni programom PRAAT (Preuzeto iz Lieshout, 2003).**

Još jedna bitna značajka ovog programa je i anotiranje govora, odnosno označavanje i segmentiranje govora, bilo da se radi o označavanju čitavih rečenica, pojedinačnih riječi ili pak glasova pomoću propisanih simbola Međunarodne udruge fonetičara (*International Phonetic Association*, IPA). Ova značajka programa ne samo da može poslužiti fonetičarima, fonolozima te logopedima, primjerice u mjerenju mucajućih slogova, pauza, produžavanja glasova kod poremećaja tečnosti govora ili čak mjerenju prosječne duljine iskaza u kontekstu jezičnog razvoja, već može imati ulogu i u području forenzičke akustike i fonetike. Ovdje valja ponovno spomenuti ulogu lingvističkih obilježja u identifikaciji govornika, gdje postaje jasno kako se pomoću ovog programa mogu na primjer bilježiti i određene specifične i

upadljive riječi koje se često pojavljuju u vokabularu osobe sa snimke. Isto tako, može se raditi o osumnjičeniku koji ima poremećaj tečnosti govora pa za cilj imamo usporediti značajke njegova govora i govora osobe sa snimke.

Nadalje, još jedna posebnost PRAAT-a nalazi se u mogućnosti programiranja, a time i automatiziranja određenih programskih naredbi. To znači da sve ono što u programu PRAAT želimo izračunati, možemo zapravo automatizirati i to prilagoditi vlastitim potrebama. Kako bi sve bilo jasnije slijedi primjer; recimo da imamo snimku govora u trajanju od jedne minute te u toj snimci želimo izračunati broj ponavljajućih slogova. Prvo ćemo odijeliti/anotirati slogove (100 slogova) na način da upotrijebimo broj 0 za označavanje onih slogova koji su pravilno izgovoreni, a broj 1 za one koji su ponavljani. Nakon što je prethodno učinjeno trebamo pobrojati koliko puta se u iskazu od 100 slogova pojavljuje oznaka 1, no sada nastupa uloga programiranja koja omogućava da se sljedeći korak odradi automatski te da PRAAT izbaci novi programski prozor iz kojega će se taj podatak moći iščitati. Iako je ovo jedan vrlo jednostavan primjer, programiranje naredbi ide znatno dalje te može biti vrlo kompleksne naravi. Prema tome, onaj tko je umješšan u stvaranju ovakvih i sličnih kodova, može samostalno automatizirati bilo koju naredbu u PRAAT-u, no već pozamašna količina gotovih naredbi može se pronaći na internetskoj stranici <http://www.fon.hum.uva.nl/praat/manual/Scripting.html>.

Boersma i van Heuven (2001) naglašavaju da ne samo da je PRAAT jedan od najcjelovitijih programa trenutno dostupnih koji je besplatno distribuiran, nego i u sebi sadržava vrlo precizne algoritme. Naime, algoritam zaslužan za analizu visine glasa/govora je kao takav najprecizniji od svih takvih algoritama unutar programa koji trenutno postoje na tržištu. No, ovisno o potrebama i karakteristikama koje se istražuju uvijek će biti onih programa koje će zbog nekih drugih značajki biti u prednosti nad PRAAT-om.

#### **1.4. Prednosti i ograničenja metoda za procjenu glasa**

Bijhold i suradnici (2007) predstavili su svoj rad na petnaestom simpoziju forenzičke znanosti INTERPOL-a koji obuhvaća literaturu u razdoblju od 2004. do 2007. godine te razlikuju dva pristupa identifikaciji govornika. Jedan od pristupa je auditivno-akustički, dok je drugi vezan uz novija saznanja u ovom području, a odnosi se na automatsku identifikaciju. Ta se dva pristupa u posljednje vrijeme sve više isprepliću, budući da stručnjaci koriste metode i

kombiniraju oba pristupa u vještačenju glasa i govora. Na trinaestom simpoziju INTERPOL-a održanom u Lyonu, Francuska, 2001. godine Broeders spominje fonetsko-akustički pristup, analitičko-akustički pristup koji je kombinirani poluautomatski te globalni automatski. Automatsko prepoznavanje govornika koristi se drugačijim metodama i obuhvaća različita obilježja unutar procjene za razliku od auditivno-akustičkog pristupa te ne podrazumijeva ulogu nekog od stručnjaka, već se ta obilježja automatski ekstrahiraju pomoću raznih matematičkih algoritama. Kao mana automatskog pristupa izdvaja se veća vjerojatnost pogreške (npr. pogreške u F0 izračunu ukoliko postoje značajne varijacije) naspram auditivno-akustičkog u kojem stručnjak sa znanjem fonetike i lingvistike takve pogreške koje proizlaze iz automatiziranih rutina može ispraviti (Bijhold i sur., 2007). Još jedna prednost auditivnog pristupa je mogućnost detaljnog izvještaja o mjerama i rezultatima procjene na temelju akademskog obrazovanja u znanostima koje se bave proučavanjem govora i jezika. S druge strane, automatskim prepoznavanjem govornika dobiva se mogućnost obrade velike količine podataka u kratkom vremenskom razdoblju, što je od koristi u slučajevima postojanja mnogo audio zapisa govora i to s većim brojem govornika. Nadalje, možda nešto još važnije je da u slučaju automatskog pristupa strani jezik ne igra ulogu, što uvijek predstavlja prepreku kod auditivno-akustičkog pristupa (Bijhold i sur., 2007).

Većina se metoda koje koriste automatske glasovne usporedbe ili sustave verifikacije govornika baziraju isključivo na obilježjima vokalnog trakta, dok su samo neki sustavi napravljeni kako bi se pokušalo uzeti u obzir i neka obilježja samog govora. No, ta dva „dijela“, glas i govor, izrazito je teško kombinirati unutar jednog sustava za analizu te je vrlo važno znati ih razlikovati upravo jer se analiziraju na drugačiji način te se time i izražavaju različitim rezultatima (Lindh, 2017).

I ljudsko uho, koje ovdje podrazumijeva slušanje i daljnju obradu signala u mozgu, ima svoje prednosti i nedostatke. Kao jedna od glavnih prednosti izdvaja se, kao što je već spomenuto, sposobnost ljudskog mozga prepoznati glas već poznatog mu govornika čak i u zahtjevnim uvjetima gdje je prisutna buka ili se radi o govorniku koji pokušava maskirati svoj prirodni glas i način govora, ili je pak pod stresom koji mijenja njegove glasovne karakteristike. Zatim, sposobnost uočavanja idiosinkratskih obilježja glasa i govora pojedinca. Na primjer, ako neka osoba kihne na određeni način koji je upadljiv, idućeg puta kada se dogodi isto ta osoba koja je to čula i prvi puta će to vrlo brzo primijetiti i zaključiti da postoji velika vjerojatnost da se radi o istoj osobi. No, isto tako u obzir treba uzeti i ljudske

nedostatke. Samo neki od njih koji će utjecati na točnost identifikacije su: utjecaj konteksta na odluku, mogućnost pogreške, duljina zadržavanja zvuka nečijeg glasa u pamćenju, pažnja te uvježbanost (Hansen i Hasan, 2015).

Lindh (2017) u jednoj od svojih studija uspoređuje razlike između automatske glasovne usporedbe i ljudske, perceptualne prosudbe glasovne sličnosti te je uočio da u prisutnosti sličnih obilježja govora između dva govornika, kao što su pauze i govorna brzina, a istovremene odsutnosti velikih razlika u perceptualnim obilježjima glasa, slušatelji postaju pristrani govornim obilježjima i svoju odluku temelje upravo na tim razlikama. Ovo ne govori u prilog tome da govorne karakteristike u ovom slučaju kontradiktornosti treba ignorirati, već da prilikom perceptualnog prosuđivanja o sličnosti govornika treba biti vrlo oprezan i u cijelosti uzeti u obzir i glasovne i govorne karakteristike, čak i u slučajevima kada se dva glasa čine prosječnima i kada se između njih ne uočava očita razlika. Također, isti je autor dobio rezultate koji govore u prilog postojanju korelacije između onoga što slušatelj percipira kao sličnost u glasu govornika i onih rezultata koje daje automatska analiza glasa. Kao što je već ranije naglašeno, trenutno ti automatski programi koji analiziraju glas još uvijek ne mogu uključiti u svoju obradu fonetske ili lingvističke informacije, stoga se govorni parametri obrađuju zasebno i to perceptualno od strane iskusnog stručnjaka čija se zapažanja kasnije „kombiniraju“ s prethodno spomenutim parametrima akustičke analize glasa (Lindh, 2017).

Auditivna analiza putem slušanja nekog govora za neke je dijelove govornog signala puno bolji izbor, to su oni dijelovi koji se mogu čuti i koje stručna osoba može objektivno opisati. To bi se primjerice odnosilo na različite načine izgovora vokala. Ipak, s druge strane postoje ona obilježja koja su matematički definirana, a to su parametri govornog signala koji se upotrebljavaju u automatskim analizama/sustavima, a to su na primjer fundamentalna frekvencija (F0) te formantska frekvencijska širina pojasa (eng. *formant frequency bandwidth*) (Hansen i Hasan, 2015).

## **1.5. Akustika govornih zvukova**

### ***1.5.1. Fundamentalna frekvencija***

Fundamentalna frekvencija (F0) je osnovna komponenta govornog zvuka; to je frekvencija kojom vibriraju glasiljke (Heđever, 2009). U muškaraca prosječna brzina titranja glasnica tijekom govora iznosi 100 – 150 Hz, u žena 180 – 220 Hz, dok kod djece ona iznosi 300 Hz (Heđever, 2012).

U pokušaju pronalaska akustičkog parametra glasa čiju će vrijednost biti lako dobiti automatskom analizom putem računalnih programa te koji će moći pripadati, odnosno referirati se na određenu populaciju ukazujući na vrijednost koja je uobičajena ili nije, dolazi se do odabira fundamentalne frekvencije (F0). Fundamentalna frekvencija ima karakteristike razlikovanja ili tipičnosti za određenu populaciju te je lako mjerljiva zbog čega se smatra jakim obilježjem u procesu identificiranja govornika (Lindh, 2017).

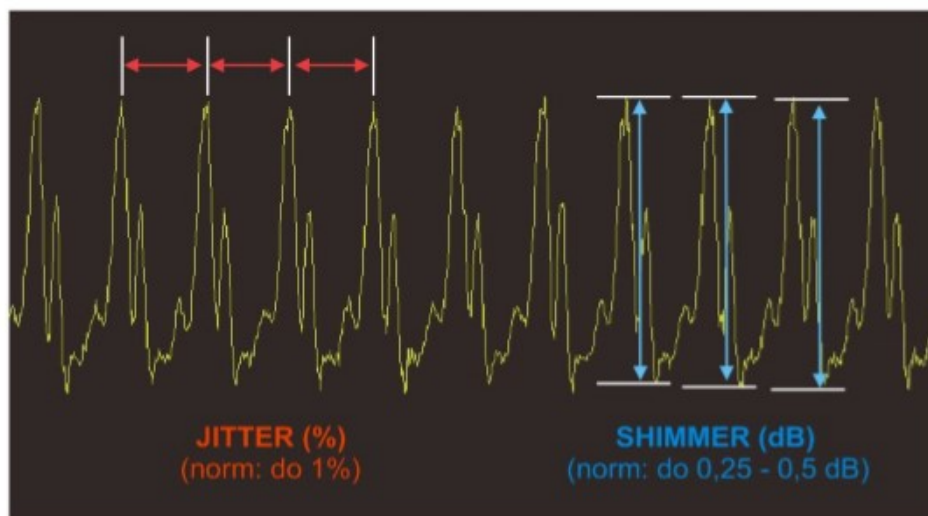
Najčešće spomenut parametar u studijama utjecaja stresa na govor je fundamentalna frekvencija ili F0. Marianne Jessen (2006) u svom radu dokazuje da dolazi do povišene vrijednosti F0 u trenucima doživljavanja psihološkog stresa pri rješavanju različitih kognitivnih zadataka uz fizičku napetost, buku ili vremensko ograničenje. Kunzel i sur. (1995) navode kako je prosječna vrijednost fundamentalne frekvencije 'najsnažniji' akustički parametar u postupku forenzičke identifikacije glasa koji se ističe kao specifičan za svakog govornika. No, kao što je već spomenuto, na temelju samo jednog značajnog parametra nije moguće te je nedopustivo donijeti odluku o govornikovom identitetu, štoviše i zbog toga što postoje značajne varijacije u vrijednosti F0 u istoga govornika kada se uspoređuju različiti uzorci snimaka glasa. Kunzel i sur. (1995) nadalje navode da bi razlog tim varijacijama mogla biti interakcija F0 s drugim akustičkim parametrima kao što su tempo, odnosno brzina govora te njegova glasnoća. Iako, u svom daljnjem radu nije došao do statistički značajne korelacije između navedenih parametara. No, treba uzeti u obzir da istraživanje nije provedeno na spontanom govoru ili na stvarnim forenzičkim uzorcima govora. Ipak, forenzički su fonetičari vjerojatno zadovoljni ovakvim podatkom, budući da im to dopušta uspoređivanje snimki govora različite govorne brzine. Drugim riječima, prije postupka usporedbe snimaka govora različite govorne brzine nije potrebno učiniti nikakve algoritamske korekcije navedenih parametara (Kunzel, 1995).

Najučestalije se za izražavanje vrijednosti fundamentalne frekvencije spominju aritmetička sredina te standardna devijacija (Rose, 2002). No, jasno je kako bi te vrijednosti mogle ovisiti o trajanju nekog uzorka glasa ili govora što predstavlja problem jer u literaturi još ne postoji usuglašenost autora oko minimalne duljine određene snimke govornika kako bi te vrijednosti bile reprezentativne, a i pouzdane (Lindh, 2017). Tako se u literaturi nailazi na različite vrijednosti te neki autori govore da je potrebna snimka od minimalno 14 sekundi (Horii, 1975), zatim drugi autor navodi da je potrebno 60 sekundi (Nolan, 1983), a postoje podaci i o minimalnom trajanju snimke od 2 minute (Baldwin i French, 1990; prema Lindh, 2017), no te brojke također mogu biti zavisne od jezika kojim se govornim služi. Braun (1995; prema Lindh, 2017) navodi kako statističke vrijednosti koje opisuju fundamentalnu frekvenciju nekog govornika ovise o psihološkim i fiziološkim čimbenicima te uz to navodi da je trajanje snimke govornika u rasponu od 15 do 20 sekundi zapravo dovoljno da te vrijednosti budu uzete kao pouzdane. No, vrijednosti  $F_0$  u govornika su obično pozitivno distribuirane (Jassem i sur., 1973; prema Lindh, 2017), što navodi na zaključak da vrijednosti  $F_0$  neće biti simetrično raspodijeljene oko aritmetičke sredine. Uzimajući u obzir i dokaze da dolazi do povećanja standardne devijacije ukoliko dolazi do varijacija živosti nečijeg govora (Traunmüller i Eriksson, 1995), odnosno življi govor automatski će povećati i vrijednost aritmetičke sredine te je jasno da postoji potreba za upotrebom druge mjere koja će biti manje podložna ovakvim varijacijama (Lindh, 2017), stoga bi možda točnije bilo upotrijebiti centralnu vrijednost.

### ***1.5.2. Jitter i Shimmer***

Termin *jitter* upotrebljava se za opis varijacija frekvencije osnovnog laringalnog tona, odnosno daje podatak o mikro nepravilnostima u brzini titranja glasnica. Takve mikro varijacije nazivaju se još i perturbacijama te njih nije moguće auditivno zamijetiti, već se one mogu uočiti i numerički prikazati akustičkom analizom oscilograma fonacije vokala. Najčešće se za normalnu vrijednost *jittera* uzima prosječno odstupanje u brzini vibriranja glasnica do 1%. Prema tome, ako uzmemo za primjer da se radi o vrijednosti  $F_0$  od 100 Hz, tada će *jitter* od 1% iznositi oko 1 Hz (Heðever, 2012). Što je njegova vrijednost manja glas će biti bolje kvalitete i čišće intonacije, no što je *jitter* veće vrijednosti glas će biti lošije kvalitete i često ga se tada može okarakterizirati kao hrapavog. No, postoje i fluktuacije amplitude koje dovode do intenzitetske nepravilnosti za koju se koristi termin *shimmer*. *Shimmer*, odnosno

brzo kolebanje amplitude izražava se decibelima, a normativne vrijednosti kreću se između 0.25 i 0.5 dB, a što je ta vrijednost veća glas se može opisati kao promukao (Heðever, 2012). Oba se ova parametra najčešće upotrebljavaju pri opisu kvalitete glasa u logopediji, no nažalost i dalje se ne koriste dovoljno učestalo u forenzičkoj identifikaciji govornika (Dalyi sur., 2016). *Shimmer* i *jitter* ne postoje u drugim zvukovima doli glasovnog govora (Heðever, 2009), ipak oni nisu dovoljni sami po sebi kao akustički parametri glasa kako bi se identificirao govornik. No, važan su podatak i vrlo dobra nadopuna spektralnoj analizi i prozodijskim obilježjima govora, a rezultati istraživanja ukazuju da su apsolutne mjere *jittera* i *shimmera* od većeg značaja kod diskriminacije govornika od njihovih relativnih mjera (Farrus i Hernando, 2009).



Slika 6: Grafički prikaz jittera (frekvencijskih perturbacija) i shimmera (amplitudnih fluktuacija) (Preuzeto iz Heðever, 2012).

### 1.5.3. Formanti

Kvistedal (2000) objašnjava akustičku funkciju vokalnih organa putem modela izvora i filtra. Model kaže da postoje dva različita izvora koja se upotrebljavaju u govoru, prvi čine vibracije glasnica, a drugi turbulencija zračne struje. Vibracijom glasnica, koja se naziva fonacijom, dolazi do periodičnih kompleksnih valova, a turbulencijom zračne struje do aperiodičkih kompleksnih valova, odnosno gibanja. Rezonantne šupljine u ovom modelu imaju ulogu filtra koji vrše „selekciju“ frekvencija prije nego dođe do krajnjeg, izlaznog „produkta“ u obliku govornih glasova. Prolaskom zračne struje kroz rezonantne šupljine neke



će frekvencije biti pojačane stvarajući tako frekvencijske vrhove (eng. *frequency peaks*) u frekvencijskom spektru koji čine formante. Promjena u frekvenciji i trajanju formanta mogu se uočiti na spektrogramskom prikazu kao linije koje se dižu i spuštaju, odnosno uočava se variranje visine formantskog traga unutar frekvencijskog opsega pojedinog formanta, a naziva se tranzicijom formanata (Heðever, 2009). U postupku identifikacije govornika analiza formanata vrši se na najmanje deset (10) do dvadeset (20) riječi s najmanje dva (2) do tri (3) ili više formanata (Heðever i Blaži, 2010). Srednje vrijednosti govornih formanata su:  $F1 = 400$  Hz,  $F2 = 1,2$  kHz,  $F3 = 2$  kHz,  $F4 = 3,9$  kHz (Heðever, 2012).

Zanimanje za frekvencije formanata unutar akustičke procjene glasa znatno se povećalo u razdoblju od 2004. do 2007. godine, što može biti potaknuto događajem iz 2002. godine kada je Žalbeni sud u Sjevernoj Irskoj iznio pravilo da se postupak procjene glasa u svrhu forenzičke identifikacije ne smije svoditi samo na slušnu/auditivnu/subjektivnu procjenu te da se u obzir moraju uzeti i mjere akustičke procjene, uključujući i frekvencije formanata, te ih priložiti kao dokaze (Bijhold i sur., 2007). Iako je upotreba glasovnog spektrograma u forenzičke svrhe naišla na kritike od strane Odbora nacionalnog vijeća za istraživanje Američke nacionalne akademije znanosti (*National Research Council committee of the American National Academy of Sciences*), ona se i dalje upotrebljava i prihvaća u vještačenju kao dokaz unutar pravosuđa u SAD-u (Bolt i sur., 1979; prema Broeders, 2001).

Lindh (2017) spominje kako su u svrhu forenzičke glasovne usporedbe tradicionalno stručnjaci kao glavnu vrstu analize upotrebljavali analizu formanata na vokalima. Već je spomenuto kako se formanti mogu objasniti kao centralni frekvencijski vrhovi koji obilježavaju pojačanje akustične energije koje nastaje kao posljedica rezonantnosti te ovisi o obilježjima supralaringealnih šupljina nekog govornika. Kako ljudi artikuliraju i percipiraju vokale uglavnom ovisi o prva tri formanta. Ne samo da formanti ovise o konfiguraciji supralaringealnih šupljina, već i o anatomiji i fiziologiji supralaringealnih artikulatora, no i drugim intra-govornikovim obilježjima govora kao što su to brzina govora, vremenska preciznost (Eriksson i Wretling, 1997), kontekst (Ingram i sur., 1996) te vokalni napor (Traunmüller i Eriksson, 2000). Još uvijek nije jasno jesu li formanti kao mjera dovoljno robustni te koje točno parametre treba izdvojiti u glasovnoj usporedbi produciranih vokala između različitih govornika (Lindh, 2017). Sve do sada znalo se koristiti manualno računanje frekvencije formanata kako bi se smanjila mogućnost pogreške, no u svom radu iz 2013. Harrison pojašnjava da upotrebom LPC (*Linear Predictive Coding*) analize uz svjesnost

moгуćnosti pogreške te uz manualnu korekciju, pogreške se mogu svesti na minimum. Posebno treba pripaziti kod računanja vrijednosti formanata ako se radi o snimci govornika učinjenoj preko telefonske linije zbog posebnih obilježja filtriranja putem telefonske linije, odnosno gubitka govornog signala koji se nalazi u govornom području do 300 Hz (Jovanović, 2009). Studija koju je Kunzel proveo 2001. godine potvrđuje značajan utjecaj ovih obilježja filtriranja glasa putem telefonske linije na vrijednosti F1 kod vokala /a/, no ne i na vrijednosti viših formanata (Kunzel, 2002). Nadalje, novija istraživanja govore u prilog upotrebi Long-Term-Formant (LTF) analize, odnosno daju prednost ovom načinu analiziranja formanata pred analizom na samo jednom izgovorenom vokalu (Gold i sur., 2013; Jessen i sur., 2013). Prema svemu navedenom očito je da još nije u potpunosti jasno koliko ove vrijednosti formanata mogu pridonijeti u razlikovanju više govornika, no sigurno je da su vrijedni podatci kada se analiziraju uz ostale parametre (Lindh, 2017).

Long-term formant distributions (LTFD) predstavlja uprosječene vrijednosti frekvencija za svakog od prva četiri formanta (F1-F4) mjerene na svim izgovorenim vokalima koje govornik izgovara na jednoj snimci. Već su prijašnje studije navele kako LTFD ima dobru diskriminatornu vrijednost kada se koristi u forenzičke svrhe za usporedbu govornika te neki autori (Becker i sur., 2008) spominju kako bi ta procedura, koja je prilično jaka, mogla postati dijelom baterije metoda koje se trenutno koriste u ovom području. Nolan i Grigoros (2005) prvi su koji govore o upotrebi LTFD metode u forenzičkoj usporedbi govornika, uz navode da LTFD daje jasnu sliku svakog od prva četiri formanta, istovremeno pružajući informacije o veličini vokalnog trakta govornika dajući maksimalnu vrijednost LTFD te o nekim govornim navikama kao zaokruživanju usana tijekom govora ili karakteristikama palatalizacije. Gold i sur. (2013) istražuju tu istu metodu te pokušavaju doći do zaključka o njezinoj korisnosti pri forenzičkoj usporedbi govornika te dobivaju rezultate kao i neki od prijašnjih autora (McDougall, 2004) koji kažu da su LTFD3 i LTFD4 više indikativni, odnosno daju više informaciju koje mogu utjecati na odluku pri identifikaciji govornika, budući da su ovisni o govornikovom ponašanju i anatomske obilježjima. Ti viši formanti, za razliku od F1 i F2 koji nose više lingvističke informacije, povezani su s rezonacijom u manjim šupljinama vokalnog trakta čineći ih tako manje indikativnima u intra-govornikovim usporedbama, a više u inter-govornikovim usporedbama (Rose, 2002). Uz ove rezultate Gold i sur. (2013) napominju kako LTFD metoda daje vrijednosti koje su neovisne o dijalektu govornika kao i o govornoj brzini. No, još zaključuju kako je pouzdanija kada se upotrebljava

za usporedbu različitih govornika (97.4% točnosti), nego prilikom uspoređivanja snimaka govora istog govornika (84% točnost).

## **1.6. Glotalni izvor glasa**

Studije obilježja glotalnog glasovnog izvora zauzimaju bitan položaj u ovom polju, ako se uzmu u obzir jednostavnost njegova izdvajanja iz spontanog ljudskog govora te dosadašnji rezultati istraživanja koji govore u prilog njihovoj diskriminatornoj značajnosti. Druga bitna značajka ovog parametra bila bi neovisnost od ostalih tradicionalnih obilježja vokalnog trakta, budući da se njegova vrijednost dobiva inverznim filtriranjem vokalnog trakta (Gómez-Vilda, i sur., 2009). San Segundo i Gomez-Vilda (2014) objavili su istraživanje učinjeno na jednojajčanim i dvojajčanim blizancima, ne blizanačkoj braći te ispitanicima koji nisu bili u srodstvu, kako bi ispitali važnost ovog parametra za područje forenzičke usporedbe govornika (usporedba glasa počinitelja i osumnjičenika). Ističu kako važnost ovakvih istraživanja glasa u blizanca leži u genetičkoj identičnosti jednojajčanih blizanaca te velikoj genetskoj sličnosti u dvojajčanih blizanaca, što uz jednake okolnosti odgoja/okoline ponekad otežava raspoznavanje njihovih glasova. Još jedan značaj ovog obilježja glotalnog izvora glasa jesu dva kriterija koja ispunjava, a to su velika intergovornikova varijabilnost te niska intra-govornikova varijabilnost. Što bi značilo da će se najveća sličnost pronaći u glasu monozigotnih blizanaca, a najmanja u usporedbi govornika koji nisu ni u nikakvom srodstvu. U svrhu svog istraživanja San Segundo i Gomez-Vilda (2014) kod svojih su ispitanika izdvojili fonaciju produženog schwa vokala u trajanju od 200 milisekundi tijekom spontanog govora snimljenog putem razgovora s ispitivačem preko telefonske linije te su ga prvo analizirali auditivno i spektrogramski pomoću programa Praat (Boersma i Weenink, 2012). Schwa vokal čini se kao bolji izbor u ovakvim studijama vezanim za forenzičku fonetiku i akustiku od produljene fonacije vokala jer je uzet iz spontanog govora u trenucima kada se osoba pokušava dosjetiti što će sljedeće reći ili se nečega ne može dosjetiti, stoga i traje dulje od drugih vokala tijekom spontanog govora. BioMet®Soft programom (2010) parametri glotalnog izvora dobiveni inverznim filtriranjem (Gómez-Vilda i sur., 2009) mogu se podijeliti u nekoliko podgrupa od kojih su neke: a) vrijednost  $f_0$ , b) kepralni koeficijent, c) koeficijent glotalnog zatvaranja te d) vrijednosti tremora. Isti je program upotrijebljen kako bi se izvršila usporedba govornika te su dobiveni rezultati naposljetku odgovarali prvotnim pretpostavkama o sličnosti i razlikama spomenutih parametara između jednojajčanih, dvojajčanih blizanaca, neblizanačke braće te govornicima

koji nisu u srodstvu, uz velike intra-govornikove varijacije u jednog ispitanika što se kasnije pokazalo posljedicom hormonalnog poremećaja uslijed hipertireoze.

### **1.7. Varijabilnost govornog signala**

Kao jedan od glavnih izazova forenzičke identifikacije govornika navodi se varijacija govornog signala. Govorni signal podložan je znatnoj razini varijacije, pa tako ako isti govornik izgovori istu riječ nekoliko puta ona neće imati jednake akustičke i spektralne vrijednosti (Eckert i Rickford, 2001). Nadalje, razni uređaji za snimanje i transmisiju govornog signala dodatno naglašavaju problem identifikacije jer ponekad otežavaju prepoznavanje čak i onih glasova govornika koji su nam dobro poznati zbog načina na koji su govorne frekvencije maskirane/filtrirane kada se signal prenosi tim uređajima. Još k tome možemo dodati i utjecaj bolesti na glasovne varijacije, ali i različit napor ili snagu prilikom govorenja, na primjer govor u šaptu ili jako vikanje (Hansen i Hasan, 2015).

Hansen i Hasan (2015) dijele izvore varijabilnosti glasa na tri šire skupine :

- 1) Ovisne o govorniku → stres, emocije, vokalni napor, govorni stil, medicinsko stanje, medikamenti, intoksikacija, maskiranje prirodnog načina govora...
- 2) Ovisne o vrsti konverzacije → razgovor između dvoje ljudi, javni nastup, monolog, razgovor putem telefona i mobilnih uređaja te putem računalnih aplikacija...
- 3) Ovisne o tehnologiji → vrsta mikrofona, okolinska/pozadinska buka, akustika u prostoriji u kojoj je govorni signal sniman, trajanje snimke, brzina uzorkovanja...

Prva skupina nabraja one uzroke koji su unutarnji/internalizirani te podrazumijeva spektar promjena koje nastaju kao posljedica različitih načina na koji neka osoba govori. Druga se pak odnosi na različite situacije u kojima ljudi komuniciraju, bilo da se radi o interakciji između ljudi uživo ili preko drugih sustava za transmisiju, ali može se odnositi i na razlike u jeziku i dijalektu. Vanjski ili eksterni uzroci nalaze se u trećoj skupini, a varijacije govornog signala u tom slučaju posljedica su mjesta gdje je govor snimam te načina na koji je isto učinjeno.

### ***1.7.1. Emocije i govor***

Kako Johnstone (2001) navodi, tijekom posljednja dva desetljeća znanstvenici u području forenzičke akustike i fonetike sve uspješnije pronalaze povezanost između određenih akustičkih parametara te govornikovih stavova i emocija temeljem pretpostavki o fiziološkim promjenama pod utjecajem istih (Banse i Scherer, 1996; prema Johnstone, 2001). Iako, i dalje nedostaje empirijskih dokaza koji dovode fiziološke promjene tijekom emocionalnih odgovora u vezu s akustičkim mjerama te istraživanja koja integrativno pristupaju ovom pitanju koristeći veći broj parametara koji opisuju glas.

Johnstone (2001) u svojoj doktorskoj disertaciji govori o pokušaju simuliranja i prikazivanja utjecaja emocija na kvalitetu glasa i to pomoću video igrice. U istraživanju je sudjelovalo 36 muških osoba, studenata Sveučilišta u Genevi, kojima je zadatak bio igrati video igru u kojoj su da bi došli do krajnjeg cilja nailazili na razne prepreke. Pozornost je bila usmjerena i na način nošenja u nepredvidivim situacijama. Sudionici su bili dužni tijekom igre u trenucima manipuliranja izjaviti kako se trenutno osjećaju tako da izgovore rečenicu „U ovom trenutku osjećam se....“ te dopunjavaju izrečeno sa jednim od sljedećeg: iziritirano, razočarano, zadovoljno, naživcirano, olakšavajuće, bespomoćno ili uznemireno. Pri tome im je bio sniman glas i potom analiziran programom LabSpeech te je upotrijebljena površinska elektromiografija kojom se mjerila mišićna aktivnost podlaktice na onoj ruci kojom se nije igrala video igra, otkucaji srca mjereni su elektrokardiogramom, a respiratorna aktivnost bilježila se kroz mjerenje naprezanja u području prsišta. Rezultati ukazuju na povišen F0 i smanjenje dubine disanja u trenucima kada im je bilo teško nositi se sa situacijom koju su opisivali kao stresnom te ubrzan rad srca prilikom nailaženja na prepreke. S druge strane, nije postojala značajna razlika u prosječnim vrijednostima akustičkog signala te mišićnoj aktivnosti podlaktice u trenucima koje su opisivali kao opuštenima ili stresnim (Johnstone i Scherer, 1999)

U ovakvim studijama koje govore o korelaciji između određenih statističkih mjera govora i emocionalnih stanja, najčešće je riječ o visini glasa, energiji, artikulaciji te spektralnom obliku. Za primjer, osjećaj tuge rezultira smanjenom standardnom devijacijom visine te sporijom brzinom govora, dok ljutnja s druge strane dovodi do povećane devijacije visine i brzine govora (Nogueiras i sur, 2001). Roberts je u svom preliminarnom istraživanju iz 2011. pokušala doznati mogu li se opaziti razlike u fundamentalnog frekvenciji (F0) i kvaliteti produkcije vokala kada se uspoređuju snimke autentičnog govora u nevolji, što je

ujedno i jaka strana ovog istraživanja, te onog simuliranog od strane glumaca. Valja naglasiti da je ovdje riječ o britanskom engleskom jeziku te da su snimke obrađene PRAAT programom za akustičku obradu glasa. Rezultati su ukazali na povišen F0 unutar obje skupine govornika, onih koji su bili snimljeni u stvarnom trenutku nevolje i onih koji su isto odglumili. Nadalje, promjene su uočene i u produkciji vokala /i/, odnosno došlo je do povišenja prvog formanta (F1) u skupini govornika glumaca, a u drugoj skupini zabilježeno je smanjenje drugog formanta (F2). Promjene su opažene i u proizvodnji /eu/ diftonga, no one nisu odvajale dvije skupine govornika. Opseg vrijednosti F0 razlikovao se i između pojedinaca unutar iste skupine govornika, što ide u prilog tome da osim univerzalne fiziološke reakcije (Jesen, 2006) kao što su napetost laringealne muskulature i glasnica (Roberts, 2011) i psihološka procjena nevolje u određenoj situaciji igra ulogu u varijacijama akustičkih parametara (Kirchhubel i sur., 2011).

Generalno govoreći, već se zna kako određena emocionalna stanja, a uz njih i neka obilježja osobnosti, utječu na glasovne karakteristike. Na primjer, anksioznost, nesigurnost i sklonost neurotizmu dovode do pretjeranog istjecanja zraka prilikom vibriranja glasnica što glas čini šumnim, žalost i tuga sužavaju raspon visine glasa i snižen intenzitet te na primjer sreća i radost utječu na glasovne karakteristike tako što dovode do frekvencijskog povišenja i intenzitetskog povećanja glasa (Heðever, 2009).

#### **1.7.1.1. Sustavi za analiziranje stresa u glasu**

Kako zapravo definirati stres i kako on utječe na govor uopće? Vrlo je teško odabrati definiciju kojom bi bile zadovoljne sve one struke koje su u dodiru s ovom problematikom, uzimajući u obzir da ljudsko biće stres može doživjeti mentalno/psihološki, emocionalno ili fizički. Općenito su definicije stresa podijeljene na one koje su usmjerene na podražaj i one koju su usmjerene na odgovor. No, za doživljaj psihološkog stresa, fizički podražaj ne mora biti prisutan, već ovisi o uvjerenjima i kognitivnom stanju osobe, što ujedno znači da je taj doživljaj pod velikim utjecajem individualnih razlika (Kirchhubel i sur., 2011).

Sustavi za analiziranje stresa u glasu govornika (engl. *Voice Stress Analysis, VSA*) računalni su sustavi koji služe za prepoznavanje laži u iskazu govornika putem mjerenja razine stresa u glasu pojedinca, odnosno promjene u obilježjima glasa uslijed različitih psiholoških/mentalnih stanja. Većina ih je u obliku računalnih softvera, a drugi se na tržištu moću naći kao elektronički uređaji s ugrađenim softverom. Prema navodima Haddad i sur. (2002) ovi sustavi imaju manje ograničenja od tehnologije poligrafa, a također su

jednostavniji za upotrebu te neinvazivni i cjenovno pristupačniji. Kako se tržište takvih računalnih sustava neprestano povećava, postavlja se pitanje učinkovitosti takvih komercijalnih sustava koji se mogu nabaviti po vrlo niskim cijenama. No, uz samu pristupačnost cijene ovih sustava, kao prednost se još navode i mogućnost da se pomoću njih analizira govor u realnom vremenu, kao i sama snimka govora. Također, konzumacija opijata, medicinska stanja te dob govornika ne utječu na rezultate analize, uređaj ne treba biti fizički vezan uz osobu te uzorak govora može biti bilo kakav input, sve od duljih iskaza do jedne riječi, čak i jecanja ili stenjanja (Haddad i sur., 2002). VSA radi na principu prepoznavanja malih oscilacija u fiziološkom mikro-tremoru vibriranja glasnica tijekom normalne kontrakcije mišića u trenutku davanja lažnog iskaza. U trenutku izdavanja iskaza, ukoliko osoba vara i ne govori istinu, zbog stresa kojeg tada doživljava mikro-tremori se smanjuju u frekvenciji i amplitudi (Heðever, 2009).

Ipak, Haddad i sur. su (2002) uspoređujući dva VSA softvera (*Diogenes Lantern System* i *Vericator System*) uočili da zapravo ne rade na tom principu te da je istina da bilježe stres u govoru pojedinca, ali samim time ne otkrivaju ujedno radi li se o lažnom iskazu. To bi značilo da upotrebom takvih računalnih programa ne bi mogli sa sigurnošću reći o kojem obliku stresa je riječ, odnosno jesu li promjene u govoru prisutne uslijed emocionalnog stresa radi same ispitne situacije ili zbog laži koje u tom trenutku osoba izgovara.

## **2. CILJ RADA**

Cilj je ovoga diplomskog rada prikazati današnji status forenzičke akustike i fonetike u području vještačenja glasa i govora kroz pregled dosadašnjih znanstvenih i stručnih radova uz prikaz mogućnosti, novina i ograničenja koje ovo područje sa sobom nosi.



### 3. OBILJEŽJA GOVORNIKA

Kao što je već spomenuto, govor je spoj glasa, lingvističkih i ne lingvističkih elemenata akustičkog signala. Lingvistički elementi obuhvaćaju glasovne jedinice kao što su alofoni, zatim više razine lingvistike kao što su morfologija, leksik i sintaksa te pragmatika. Dok se ne lingvistički elementi odnose na razne zastoje u govoru, oklijevanja te neke druge zvukove koje ljudi proizvode tijekom govora, na primjer kašalj (Lindh, 2017).

Kovač (2012) ukazuje da bi postupak identifikacije govornika trebao procjenjivati ona govorna obilježja koja se smatraju specifičnima za pojedinu osobu te navodi da u tu kategoriju pripadaju visina i boja glasa, naglasak, dijalekt, brzina i prekidi u govoru, artikulacija ili način izgovaranja glasova, kvaliteta glasa, poštapalice i moguće nepravilnosti u govoru.

Osim samih akustičkih parametara, Jessen (2007) spominje šest domena na koje se obično obraća pozornost prilikom forenzičke analize govora: dijalekt, strani naglasak, sociolekt, dob, spol i zdravstveno stanje unutar kojeg spadaju i poremećaji glasa/govora. Njemačka Federalna Policija (*njem. Bundeskriminalamt, BKA*) navodi termine „glasovna analiza“ (*eng. voice analysis*) i „glasovna usporedba“ (*eng. voice comparison*) kao dva glavna procesa koja se provode u forenzičkoj identifikaciji govornika, a razlika između tih dvaju procesa temelji se na dostupnosti materijala, odnosno snimki govora te postojanju osumnjičenika. Ako postoji pristup audio zapisima govora počinitelja, ali ne i osumnjičenik ili osuđenik, tada se isprva radi „profiliranje govornika“ koje podrazumijeva podatke poput dobi govornika, spola, jezične kompetencije te socijalne pozadine. U suprotnom, ako osumnjičeni ili optuženi već postoji tada se prelazi na glasovnu usporedbu putem auditivne procjene u kombinaciji s analizom digitalnog audio signala u skladu s najnovijom tehnologijom. Ovaj postupak se nastavlja dalje i izrazito je kompleksan, ovisno o ostalim uvjetima vezanima uz snimku i govornika, no za potrebe ovog rada bilo bi preopširno približiti isti, stoga će se u dijelu koji slijedi pisati o prethodno spomenutih šest domena u forenzičkoj procjeni govora.

#### 1.1. Spol

Prosječna fundamentalna frekvencija govornika (F0) ili osnovni laringealni ton, tj. visina glasa razlikuje se između muškaraca i žena najčešće dovoljno da uho može prepoznati spol govornika (Jessen, 2007). U odraslih govornika prosječna fundamentalna frekvencija

znatno je viša u žena te se najčešće govori o prosječnoj fundamentalnoj frekvenciji od 200 Hz u žena, a oko 120 Hz u muškaraca (Traunmuller i Eriksson, 1994; Hunter i Titze, 2010; Pepiot, 2014). Iako se smatra da je određivanje spola govornika vrlo jednostavan zadatak, što uglavnom i je tako, ali ponekad ga neki drugi uvjeti mogu učiniti izazovnim. Za primjer, u slučajevima kada govornik govori falsetnim glasom, kada je prisutan šapat ili neki drugi oblik maskiranja glasa, a isto tako ako jednostavno na temelju visine nečijeg glasa ne možemo odrediti spol jer je visina glasa „neutralna“, odnosno F0 ne daje dovoljno jasne informacije. Nadalje, može biti da zbog promjena u glasu koje su uzrokovane hormonalnim promjenama (patološkim ili nepatološkim) ili drugim medicinskim stanjima nije moguće sa sigurnošću odrediti spol govornika. Anatomske razlike između muškaraca i žena imaju utjecaj i na frekvenciju formanta koja je viša u žena (Jessen, 2007), no te se vrijednosti također razlikuju i između različitih jezika, odnosno podliježu nekim kulturalnim utjecajima (Johnson, 2006; prema Jessen, 2007). Osim ovih paralingvističkih razlika koje se mogu uočiti između muškaraca i žena, razlike su prisutne i na razini lingvistike, odnosno morfosintakse, leksika te pragmatike, a te su razlike također pod znatnim sociokulturalnim utjecajima (Jessen, 2007).

## **1.2. Dob**

U forenzičkoj se fonetici i akustici spominju dva pojma vezana uz određivanje dobi govornika. Razlikuju se kronološka (kalendarska) dob, koja je određena datumom rođenja osobe te biološka dob koja odražava starost određenih organa te fiziološko stanje govornog mehanizma koji može ubrzano starjeti pod utjecajem zlouporabe vokalnog aparata (stres, alkohol, pušenje, glasan govor). Kronološka i biološka se dob pretjerano ne razlikuju kada se radi o osobi kronološke dobi u razmaku od 20 do 25 godina, no sa starenjem ta se razlika može povećati (Linville 2001; prema Jessen, 2007). Iako su stručnjaci nešto bolji u auditivnom procjenjivanju starosti govornika od neutreniranih osoba, krajnju odluku o dobi ne bi trebalo donijeti bez dodatnih akustičkih analiza. No, upotrebom auditivne i akustičke procjene i dalje se ne može odrediti točna dob govornika, već se radi o približnoj dobi u razmaku od 10 do 20 godina. Stoga se često može naći u izvještajima kako stoji da je govornik u dobi od 25 do 40 godina. Jednako kao i u slučaju određivanja spola i starosna dob ima utjecaj na razlike u jeziku, najviše leksiku govornika (Jessen, 2007).

### **1.3. Dijalekt**

Prepoznati dijalekt kojim neka osoba govori ujedno znači odrediti područje na kojem je ta osoba živjela većinu svog života ili do kraja perioda usvajanja jezika, nakon kojeg više ne dolazi do velikih dijalektalnih promjena u slučaju preseljenja u neku drugu regiju (Jessen, 2007). Posebnosti nekog dijalekta mogu biti od velike važnosti za identifikaciju ili eliminaciju govornika (Heðever, 2009). Dva glasa koja se uspoređuju mogu biti vrlo slična ukoliko ih se samo percipira auditivno pa čak i ako se njihove akustičke vrijednosti izračunaju putem automatskih računalnih programa, dok se zapravo može raditi o dva znatno različita naglasaka ili čak jezika. Kada kažemo govornik, misli se na skup svih lingvističkih, socijalnih, pragmatičkih, bihevioralnih i idiosinkratskih karakteristika koje su prezentirane putem i glasa i govora (Lindh, 2017). Samo u hrvatskom jeziku postoje tri (3) narječja, svaki od kojih se dijeli na nekoliko dijalekata, što dovodi do krajnje brojke od ukupno 27 dijalekata. Jok k tome, hrvatski, bosanski, srpski i crnogorski, sve su jezici koji imaju sličan fonološki sustav što govornicima tih jezika omogućava nesmetanu komunikaciju (Heðever, 2009).

### **1.4. Strani naglasak**

Prisjetimo se opet Kodeksa pod točkom 6(a) „Članovi trebaju postupati s posebnim oprezom, ako izvede forenzičku analizu bilo kakve vrste na snimkama koje sadrže govor na jeziku koji nije njihov materinski jezik“ (IAFPA, 2004; preuzeto s <http://www.iafpa.net/code.htm>), dakle ukoliko je moguće vještak bi trebao izbjegavati vještačenja govornika drugog jezika. Schiller i Koster (1996) su u svojem istraživanju došli do zaključka da je pouzdanost ispravne identifikacije govornika znatno manja u slučajevima pokušaja identifikacije/eliminacije govornika stranog jezika, a pogotovo ako je riječ o jeziku koji pripada nekoj od drugih skupina jezika, odnosno koji je u manjoj mjeri povezan s materinskim jezikom onoga koji procjenjuje govor. Ponekad se ne mora raditi o stranom jeziku, ali može biti prisutan strani naglasak koji može biti manje ili više upadljiv. Prethodno spomenut dijalekt i naglasak se preklapaju i u govoru koriste naizmjenično te iako se često zamjenjuju, upućuju na različita jezična ponašanja (Heðever, 2009). Sposobnost prepoznavanja glasova ovisi o iskustvu i upoznatosti s fonološkim sustavom određenog jezika. Mada ljudi imaju mogućnost prepoznavanja glasova onih osoba koje govore stranim jezikom, to ipak čine u mnogo manjoj mjeri točno (Perrachione, Del Tufo i Gabrieli, 2011).

## **1.5. Sociolekt**

U radu forenzičara sociolekt predstavlja važnu komponentu pri donošenju zaključaka o identitetu govornika. Jezik kojim se osoba služi može putem sociolingvističkih varijabli otkrivati njezin socijalni status, razinu edukacije, profesiju kojom se bavi i sl. Leksička i sintaktička složenost kao i stil kojim osoba govori mogu poslužiti kao dobar trag k pretpostavci o razini edukacije, a leksik ujedno može odražavati govornikovu profesiju (Jessen, 2007).

## **1.6. Medicinsko stanje**

Prema Tanneru (2007) komunikacija je jedna od najviših mentalnih, fizičkih i neuroloških sposobnosti koju ljudi posjeduju te postoji znatan broj stanja, bolesti i poremećaja koji na komunikaciju mogu utjecati. Određena neurološka oboljenja mogu dovesti do toga da način komunikacije tih osoba nalikuje na onu osobe pod utjecajem intoksikacije. Ta se oboljenja generalno mogu razgraničiti na tri skupine od kojih prva čini stanja nakon moždanog udara i traumatskog oštećenja mozga, zatim neurodegenerativne bolesti kao što su to Parkinsonova bolest, multipla skleroza ili amiotrofična lateralna skleroza te posljednja kategorija koja obuhvaća demencije (Heðever, 2009). Odras navedenih stanja na komunikaciju može značajno varirati ovisno o vrsti i stupnju uznapredovalosti bolesti. No, i motorički govorni poremećaji poput govorne apraksije ili dizartrije mogu dovesti do pogrešnog zaključka da se radi o govorniku pod utjecajem intoksikacije.

## 4. ULOGA LOGOPEDA

Uloga logopeda kao sudskog vještaka u području vještačenja glasa i govora sve je samo ne beznačajna. Kao stručnjak koji se između ostalog bavi poremećajima govora, jezika, artikulacije, glasa, zatim poremećajima jezika uzrokovanim raznim neurološkim poremećajima i neurodegenerativnim bolestima te (re)habilitacijom slušanja i govora osoba s oštećenjem sluha, zasigurno može biti od značajne koristi tijekom vještačenja glasa i govora. Specifična obilježja glasa, govora i jezika prisutna su kod pojedinih poremećaja koje stručnjak bez teorijskog i praktičnog iskustva u tom području neće tako lako zamijetiti ili im pridavati preveliku pažnju, a u najgorem slučaju može pomisliti da se radi o stanju intoksikacije govornika, dok je zapravo riječ o jednom od navedenih poremećaja ili obratno. Također, uzimanje iskaza osumnjičene osobe s ciljem glasovne usporedbe između glasa osumnjičenika i osobe sa snimke može biti nezgodno za stručnjaka bez navedenog iskustva ukoliko se radi o osobi s na primjer neurološkim oboljenjima. „Produkcija zvučnog govora podrazumijeva neuromuskularne aktivnosti kojima upravlja središnji živčani sustav (SŽS)“ (Heđever, 2009; str. 64). Intoksikansi na primjer mogu imati utjecaj na motoričko govorno programiranje, motoričku govornu koordinaciju i automatski motorički govor, isto tako i na upotrebu jezika te na pamćenje i orijentaciju (Heđever, 2009). Neka neurološka oboljenja imaju simptome vrlo slične onima koji se javljaju u osoba pod utjecajem intoksikacije. Govornici a afazijom te oni koji su pretrpjeli traumatsko oštećenje mozga mogu u različitoj mjeri imati teškoća s govorom, jezikom, pamćenjem i orijentacijom. Narušena artikulacija i brzina govora te prisutnost patološkog tremora mogu biti prisutni kod osoba s Parkinsonovom bolesti, multiplom sklerozom i amiotrofičnom lateralnom sklerozom, a u kasnijem stadiju bolesti i ovisno o stupnju neurološkog oštećenja mogu nalikovati na govor pod utjecajem intoksikanata (Heđever, 2009). Zatim, bolesti vezane uz treću životnu dob, kao što su Alzheimerova bolest i drugi oblici demencije, također se u kasnijem stadiju kada dolazi do narušenog govora, jezika i pamćenja mogu rezultirati istom pogrešnom odlukom da se radi o intoksiciranoj osobi (Heđever, 2009). Prepoznati da je riječ o neurološkom oboljenju može biti vrlo važna stavka u identifikaciji govornika i donošenju krajnje odluke. No, uz ove prethodno navedene stavke, uloga logopeda značajna je i kod usporedbe govornika kod kojih su prisutni razni poremećaji svih aspekata govorne produkcije, artikulacije, fonacije, rezonancije, prozodije pa i respiracije, poremećaji glasa, poremećaji tečnosti ili mogu na temelju obilježja govora i jezika prepoznati da se radi o osobi s oštećenjem sluha.

## 5. NOVOSTI I BUDUĆA ISTRAŽIVANJA

2010. godine pri Nacionalnom institutu za standarde i tehnologiju (eng. *National Institute of Standards and Technology, NIST*) događa se bitan korak u približavanju dvije spomenute analize, auditivne/perceptivne analize glasovnih i govornih karakteristika te automatske analize glasovnih parametara (*Human Assisted Speaker Recognition, HASR*) (Greenberg i sur., 2010). Također, *Massachusetts Institute of Technology* (MIT) i *United States Secret Service* (USSS) zajedničkim snagama razvijaju fonetički u kombinaciji s automatskim sustavom za analizu glasa pod nazivom *Super Phonetic Annotation and Analysis Tool (SPAAT)* (Lindh, 2017).

Berisha i suradnici (2013) predlažu novi algoritam koji bi trebao u snimci govora osobe s dizartrijom prepoznati ona obilježja govora koja su specifična za tu govorno-glasovnu patologiju, a minimalizirati ona koja se odnose općenito na glasovne specifičnosti pojedinih govornika. Odnosno, taj bi algoritam stoga izdvajao ona obilježja koja odražavaju razlike među govornicima s dizartrijom, a ne ona obilježja koja su prisutna zbog različitih govornih stilova. Time bi se došlo do akustičkih parametara koji bi dobro diskriminirali različite govornike s dizartrijom, a koja istovremeno minimalno razlikuju zdrave govornike bez takvih teškoća. Nadalje, uočili su da se varijacije u spektralnoj energiji mogu povezati s nekim aspektima takvih govornih teškoća, dok se mjere visine glasa češće vežu uz razlike u govornom stilu. Brzina, ritam i motorička kontrola govora tri su glavna obilježja koje ovaj algoritam ujedinjuje uzimajući u obzir da su promjene u ovim obilježjima govora najčešće auditivno uočljiva u govornika s dizartrijom, a isto tako ne postoji jedna akustička mjera kojom bi se njihove vrijednosti zajednički izrazile. Time bi se izdvajanjem ovih triju obilježja dobio 'digitalni potpis' govornika s dizartrijom koji bi predstavljaju individualni opis trenutnog stanja govorno-glasovne patologije (Berisha i sur., 2013).

Zanimljiva se studija Hughes i sur. objavljuje 2016. godine u kojoj su istraživali vrijednost podataka koji se mogu dobiti iz analiziranja ispunjenih pauza tijekom spontanog govora u svrhu identifikacije govornika. Željeli su znati mogu li te ispunjene pauze (npr. um, uh) koje se pojavljuju u gotovo svih govornika poslužiti kao valjani parametar u dokazivanju identiteta govornika, odnosno u forenzičkoj glasovnoj usporedbi. Temeljem raznih dinamičkih izračuna središnjih frekvencija prvih triju formanta te trajanja vokalnosti i nazalnosti pri fonaciji „uh“ i „um“ podatci su ukazali da ovakav pristup ima izvrstan

potencijal u forenzičkoj glasovnoj usporedbi te da bi u budućnosti ovakve analize ispunjenih pauza mogle poslužiti kao nove varijable u postupku identifikacije govornika.

U procesu prepoznavanja govornika sa snimke jednim dijelom se pokušava imitirati sposobnost ljudskog mozga za obradom auditivnog signala. Iako su potrebna brojna istraživanja kako bi razumjeli način na koji naš mozak zapravo obrađuje informacije, no ono što zasigurno znamo jest da je osjetljiv na vokalni podražaj (Belin i sur., 2000), da svojom aktivacijom pokazuje kako poznaje fonologiju određenog jezika (Perrachione i sur., 2011) te da je odgovoran i za složenu temporalnu obradu zvuka (Zatorre i Belin, 2001). Nadalje, naš mozak posjeduje različite mehanizme obrade onih glasova koji su nam već poznati i onih koji to nisu, odnosno onih glasova koji su specifični za neki drugi jezik s kojim nemamo iskustva. Prema tome, bilo bi interesantno otkriti kako to mozak čini, kako sprema informacije o govornikovom identitetu, što bi potencijalno moglo poslužiti u stvaranju novih, naprednih algoritama koji bi se koristili u području automatskog prepoznavanja govornika.

Hansen i Hasan (2015) naglašavaju kako će za stručnjake pokušaj pronalaska reprezentativnih parametara koji naglašavaju identitet govornika, a uz mogućnost da se istovremeno potisnu oni koji za to nisu relevantni, uvijek biti izazov.

## 6. FORENZIČKA AKUSTIKA I FONETIKA U HRVATSKOJ

Na diplomskom studiju logopedije pri Edukacijsko-rehabilitacijskom fakultetu Sveučilišta u Zagrebu izvodi se kolegij pod nazivom “Uvod u forenzičku fonetiku i akustiku” pod nositeljstvom prof. dr. sc. Mladena Heđevera. Prema dosadašnjim saznanjima ovo je prvi kolegij takvog tipa koji se izvodi u Hrvatskoj od akademske godine 2008/09 (Heđever, 2009). Iako broj kaznenih djela za čije je razrješenje potrebno vještačenje glasa i govora raste već niz godina, kako u ostatku svijeta, tako i na području Hrvatske i regije, samo nekolicina osoba bavi se takvim vještačenjem u našoj zemlji. Jedan od njih je i prof. dr. sc. Mladen Heđever koji posljednjih desetak godina obavlja sudska vještačenja iz forenzičke fonetike i akustike te ih je do sada učinio desetak, a uključivala su procjenu autentičnosti snimke, čišćenje snimki od šumova i radi poboljšanja razumljivosti govora te naravno prepoznavanje govornika temeljem analize glasa, govora i jezika.

Također treba spomenuti prof. dr. sc. Gordanu Varošaneć Škarić, profesoricu sa Odsjeka za fonetiku, Filozofskog fakulteta Sveučilišta u Zagrebu koja se bavi forenzičkom fonetikom te je u tom području vodila jedan znanstvenoistraživački projekt, objavila nekoliko članaka te sudjelovala u nekoliko vještačenja. (<https://tkojetko.irb.hr/znanstvenikDetalji.php?sifznan=4428&podaci=projekti&sifprojekat=4882>; [http://www.ffzg.unizg.hr/fonet/?page\\_id=122](http://www.ffzg.unizg.hr/fonet/?page_id=122); <https://bib.irb.hr/lista-radova?autor=153544>).

No, i dalje ne postoje stalni sudski vještaci u ovom području, niti specijalizirana ustanova pa tako ni propisane norme i procedure. Prema tome, obavljanje ovoga posla u skladu je s ranije spomenutim Kodeksom Međunarodnog udruženja forenzičke fonetike i akustike (*The International Association for Forensic Phonetics and Acoustics* (IAFPA)) te prema pravilima Američkog odbora za vještačenja – standardi za provođenje analize glasa i govora (*American Board of Recorded Evidence – Voice Comparison Standards*) (<http://www.taracentar.hr/o-nama/forenzicna-fonetika-i-akustika/>). Također, prema već spomenutim međunarodnim standardima, da bi se neki stručnjak mogao baviti ovom problematikom mora imati pozamašno znanje i iskustvo te biti akademski obrazovan u području fonetike, govora ili patologije glasa uz vrlo dobro poznavanje znanstvenih metoda procjene i ispitivanja glasa, govora pa i jezika. Premda je u Hrvatskoj ovo područje još nedovoljno razvijeno, na studiju logopedije dobiva se dobra podloga za nastavak u smjeru te



problematike za onoga tko se u tom području pronalazi. Naime, uz spomenuti kolegij koji daje uvid u osnove forenzičke akustike i fonetike, studenti logopedije i tijekom preddiplomskog studija slušaju kao obavezni kolegij “Osnove fiziološke i govorne akustike” te mogu izabrati kolegij “Psihoakustika”, a tijekom zadnjeg semestra diplomskog studija izvodi se i kolegij “Govorna akustika” (<http://www.erf.unizg.hr/hr/studiji/preddiplomski-logopedija>). Svi ti kolegiji zajedno pružaju znatnu količinu znanja te ako ubrojimo iskustvo i znanje o govorno-jezičnoj i glasovnoj patologiji koja se stječe na ovom studiju, jasno je da su to stručnjaci koji bi nakon nekoliko godina iskustva zadovoljavali međunarodne standarde koji nalažu već navedene preduvjete za obavljanje poslova vještačenja glasa, govora i jezika.

## 7. ZAKLJUČAK

Kao što je na samom početku ovoga rada spomenuto, forenzička akustika i fonetika relativno je novo interdisciplinarno područje te će prostora za napredak uvijek postojati, još štoviše zbog međusobnog ispreplitanja nekoliko znanstvenih područja, ali i očigledno zbog napretka tehnologije koja ne samo da pomaže pri razrješavanju kaznenih djela, već i u njihovom počinjavanju. Identifikacija govornika nije potpuna novost današnjice te njezini početci sežu daleko u povijest, no točnost njezina ishoda te metode koje do toga dovode neusporedivo su

napredovale. Krenuvši od same subjektivne procjene temeljem ljudskih osjetila stručnjaci sada u vještačenju glasa i govora koriste subjektivne/auditivne i objektivne/akustičke metode koje se međusobno nadopunjuju i obje daju relevantne podatke u sudskim postupcima, dok se računalni programi nastavljaju usavršavati sve preciznijim matematičkim algoritmima za izračun glasovnih parametara. Identifikacija govornika je jedan vrlo kompleksan postupak koji sadrži niz procedura koje vode prema konačnoj odluci o identitetu govornika, odnosno odgovoru u obliku vjerojatnosti te se odvija uz propisana pravila Međunarodnog udruženja forenzične fonetike i akustike (IAFPA). Budući da je govor dinamičan čin pod utjecajem je raznih čimbenika koji mogu biti unutarnji i okolinski, a svi oni čine glas i govor pojedinca varijabilnim. Stres, emocije, dob, spol, dijalekt, profesija, samo su jedni od čimbenika koji naš glas čine prepoznatljivima. No, kako je govor pod vodstvom središnjem živčanog sustava, uz navedene čimbenike varijabilnosti postoje i razni poremećaji glasa, govora i jezika koji zahtijevaju znanje i iskustvo posebnih stručnjaka u vještačenju glasa i govora. Zbog obilježja tih poremećaja može doći do zabune, a potom i krive pretpostavke o medicinskom stanju govornika sa zvučnog zapisa, u tom slučaju logoped kao stručnjak u vještačenju glasa i govora može zauzeti bitnu ulogu. Također, poznavanje fonetike, lingvistike i akustike neizostavno je u ovim postupcima. Stručnjaci logopedi na području hrvatske dobivaju kvalitetnu teorijsku i praktičnu podlogu koja može biti temelj kasnijoj gradnji znanja u području forenzičke akustike i fonetike ukoliko se odluče ići tim putem.

## 8. LITERATURA

1. Becker, T., Jessen, M., Grigoras, C. (2008): Forensic Speaker Verification Using Formant Features and Gaussian Mixture Models. INTERSPEECH 2008, 9th Annual Conference of the International Speech Communication Association, Brisbane, Australia.
2. Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B. (2000): Voice-selective areas in human auditory cortex. *Nature*, 403, 309-311.
3. Berisha, V., Sandoval, S., Utianski R., Liss, J. i Spanias, A. (2013): Selecting disorder-specific features for speech pathology fingerprinting. In: International Conference on Acoustics, Speech and Signal Processing (ICASSP).
4. Bijhold, J., Ruifrok, I., Jessen, M. i Alberink, I. (2007): Forensic audio and visual evidence 2004-2007: A review. In: 15th INTERPOL Forensic Science Symposium.
5. Boersma, P. i van Heuven, V. (2001): Speak and unSpeak with PRAAT. *Glott International*, 5(9), 341-347.
6. Bolt, R.H., Cooper, F.S., David, E.E., Denes, P.B., Pickett, J.M. i Stevens, K.N. (1973): Speaker identification by speech spectrograms: some further observations. *Journal of the Acoustical Society of America*, 5(2), 531-534.
7. Broeders, T. (2001): Forensic Speech and Audio Analysis Forensic Linguistics 1998-2001. In: 13th INTERPOL Forensic Science Symposium, Lyon, France.
8. Committee on Evaluation of Sound Spectrograms (1979): On the Theory and Practice of Voice Identification. Washington, D.C.: National Academy of Sciences.
9. Daly, I., Novotny, M., Hajaiej, Z., Garsallah, A. (2016): Accuracy of Jitter and Shimmer Measurements for Speaker in the Database TIMIT and NTIMIT. In: 10th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP).
10. DeCasper, A.J. i Spence, M.J. (1986): Prenatal Maternal Speech Influences Newborns' Perception of Speech Sound. *Infant behavior and development*, 9, 133-150.
11. Eckert, P., Rickford, J.R. (2001): *Style and Sociolinguistic Variation*. Cambridge, UK: Cambridge University Press.

12. Eriksson, A., Wretling, P. (1997): How flexible is the human voice? - A case study of mimicry. In: Fifth European Conference on Speech Communication and Technology, EUROSPEECH.
13. Eriksson, A. (2005): Tutorial on Forensic Speech Science. Part I: Forensic Phonetics. Department of Linguistics, Gothenburg University, Gothenburg, Sweden.
14. Farrus, M., Hernando, J. (2009): Using Jitter and Shimmer in speaker verification. IET Signal Process, 3(4), 247–257.
15. Gold, E., French, P. i Harrison, P. (2013): Examining long-term formant distributions as a discriminant in forensic speaker comparisons under a likelihood ratio framework. In: Proceedings of Meetings on Acoustics, 19, Montréal, Canada.
16. Gomez-Vilda, P., Fernandez-Baillo, R., Rodellar-Biarge, V., Lluís, V.N., Alvarez-Marquina, A. i sur. (2009): Glottal Source Biometrical Signature for Voice Pathology 'Detection. Speech Communication, 51(9), 759-781.
17. Greenberg, A., Martin, A., Brandschain, L., Campbell, J., Cieri, C., Doddington, G., Godfrey, J. (2010): Human Assisted Speaker Recognition. In: NIST SRE10, The Speaker and Language Recognition Workshop, Brno, Czech Republic.
18. Haddad, D., Walter, S., Ratley, R. i Smith, M. (2002): Investigation and Evaluation of Voice Stress Analysis Technology. U.S. Department of Justice final report.
19. Hansen, J.H.L., Hasan, T. (2015): Speaker recognition by machines and humans: A tutorial review. IEEE signal processing magazine, 32(6), 74–99.
20. Harrison, P.T. (2013): Making Accurate Formant Measurements: An Empirical Investigation of the Influence of the Measurement Tool, Analysis Settings and Speaker on Formant Measurements. Doctoral disertation. University of York Language and Linguistic Science.
21. Heđever, M. (2009): Uvod u forenzičnu fonetiku i akustiku. Skripta iz kolegija Forenzična akustika i fonetika na Edukacijsko-rehabilitacijskom fakultetu u Zagrebu.
22. Heđever, M, Blaži, D. (2010): Multidimenzionalna dijagnostika glasovnih i govorno-jezičnih poremećaja u forenzici. Svjetski Dan glasa. Edukacijsko-rehabilitacijski fakultet Sveučilišta u Zagrebu.
23. Heđever, M. (2010): Govorna akustika. Skripta iz kolegija Govorna akustika na Edukacijsko-rehabilitacijskom fakultetu u Zagrebu.
24. Heđever, M. (2012): Osnove fiziološke i govorne akustike. Skripta iz kolegija Osnove fiziološke i govorne akustike na Edukacijsko-rehabilitacijskom fakultetu u Zagrebu.

25. Hollien, H.F. (1990): *The Acoustics of Crime: The New Science of Forensic Phonetics*. New York: Springer.
26. Hollien, H.F. (2002): *Forensic Voice Identification*. London: Academic press.
27. Horii, Y. (1975): Some Statistical Characteristics of Voice Fundamental Frequency. *Journal of Speech, Language, and Hearing Research*, 18, 192-201.
28. Hughes, V., Wood, S. i Foulkes, P. (2016): Strength of forensic voice comparison evidence from the acoustics of filled pauses. *International Journal of Speech, Language & the Law*, 23(1), 99-132.
29. Hunter, E.J., Titze, I.R. (2010): Variations in Intensity, Fundamental Frequency, and Voicing for Teachers in Occupational Versus Non-Occupational Settings. *Journal of Speech, Language, and Hearing Research*, 53(4), 862-875.
30. Ingram, J.C.L., Prandolini, R., Ong, S. (1996): Formant trajectories as indices of phonetic variation for speaker identification. *Forensic Linguistics*, 3, 129–145.
31. Jessen, M. (2006): *Einfluss von Stress auf Sprache und Stimme: Unter besonderer Berücksichtigung polizeidienstlicher Anforderungen*. Idstein: Schulz-Kirchner Verlag.
32. Jessen, M. (2007): *Speaker Classification in Forensic Phonetics and Acoustics*. U: C. Muller (ur.) *Speaker Classification I. Lecture Notes in Computer Science* (vol. 4343, str. 180-204). Berlin, Heidelberg: Springer.
33. Jessen, M., Enzinger, E., & Jessen, M. (2013): Experiments on long-term formant analysis with gaussian mixture modeling using vocalise. In: *Conference of the International Association of Forensic Phonetics and Acoustics*, 2013.
34. Johnstone, T., Scherer, K.R., (1999): The effects of emotions on voice quality. U: Ohala, J.J., Hasegawa, Y., Ohala, M., Granville, D., Bailey, A.C. (ur.), *Proceedings of the XIVth International Congress of Phonetic Sciences*. Department of Linguistics, Univ. of California at Berkeley, (str. 2029–2032).
35. Johnstone, T. (2001): *The effect of emotion on voice production and speech acoustics*. Doctoral disertation. The University of Western Australia: Psychology Department.
36. Jovanović, N. (2009): *Uticaj načina upotrebe mobilnog telefona na dugovremeni usrednjeni spektar govora*. 17. Telekomunikacioni forum TELFOR. Diplomski rad. Elektrotehnički fakultet u Beogradu.
37. KayPENTAX, (2008): *Software instruction manual: Multi-Dimensional Voice Program (MDVP) Model 5105*.
38. Kersta, L.G. (1962): Voiceprint Identification. *Nature*, 196(4861), 1253-1257.

39. Kirchhubel, C., Howard, D.M. i Stedmon, A.W. (2011): Acoustic correlates of speech when under stress: Research, methods and future directions. *The international journal of speech, language and the law*, 18(1), 75-98.
40. Koenig, B.E. (1986): Spectrographic voice identification: A forensic survey. *The Journal of the Acoustical Society of America*, 79(6), 2088-2090.
41. Kovač, Z. (2012): Prepoznavanje osobe u forenzici na temelju glasa. *Policija i sigurnost*, 21(2), 348-356.
42. Künzel, H.J., Masthoff, H.R., Koster, J.P. (1995): The relation between speech tempo, loudness, and fundamental frequency: an important issue in forensic speaker recognition. *Science & Justice*, 35(4), 291-295.
43. Künzel, H. J. (2001): Beware of the “telephone effect”: The influence of telephone transmission on the measurement of formant frequencies. *Forensic Linguistics*, 8(1), 80-99.
44. Künzel, H. J. (2002): Rejoinder to Francis Nolan’s “The ”telephone effect“ on formants”: A response. *Forensic Linguistics*, 9(1), 83-86.
45. Kvistedal, Y.A. (2000): Forensic Voice Identification. A Research Paper in Forensic Science The University of Auckland, New Zealand.
46. Lindh, J. (2017): Forensic Comparison of Voices, Speech and Speakers: Tools and Methods in Forensic Phonetics. Doctoral disertation. Department of Philosophy, Linguistics and Theory of Science University of Gothenburg.
47. McDougall, K. (2004): Speaker-specific formant dynamics: An experiment on Australian English /aɪ/. *International Journal of Speech, Language and the Law*, 11, 103-130.
48. McGehee, F. (1937): The Reliability of the Identification of the Human Voice. *The Journal of General Psychology*, 17(2), 249-271.
49. Meuwly, D., Drygajlo, A. (2001): Forensic speaker recognition based on a Bayesian framework and Gaussian Mixture Modelling (GMM). *A Speaker Odyssey—The Speaker Recognition Workshop*, Crete, Greece, 145-150.
50. Morrison, G.S. (2009): Forensic voice comparison and the paradigm shift. *Science and Justice*, 49, 298-308.
51. Neyman, J. i Pearson, E.S. (1933): On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 231, 289-337.

52. Nicastrì, M., Ciarella, G., Gallo, L.V., Catalano, M. i Cassandro, E. (2004): Multidimensional Voice Program (MDVP) and amplitude variation parameters in euphonic adult subjects: Normative study. *Acta Otorhinolaryngologica Italica*, 24, 337-341.
53. Nogueiras, A., Moreno, A., Bonafonte, A., i Marino, J.B. (2001): Speech Emotion Recognition Using Hidden Markov Models. U: Eurospeech. Aalborg, Denmark, str. 2679-2682.
54. Nolan, F. (1983): *The Phonetic Bases of Speaker Recognition*. Cambridge: Cambridge University Press.
55. Nolan, F., i Grigoras, C. (2005): A case for formant analysis in forensic speaker identification. *The International Journal of Speech, Language and the Law*, 12, 144–173.
56. Pepiot, E. (2014): Male and female speech: a study of mean f0, f0 range, phonation type and speech rate in Parisian French and American English speakers. *Speech Prosody*, 7, 305-309.
57. Perrachione, T.K., Del Tufo, S.N., Gabrieli, J.D.E. (2011): Human voice recognition depends on language ability. *Science*, 333, 595.
58. Roberts, L. (2011): Acoustic effects of authentic and acted distress on fundamental frequency and vowel quality. *Proceedings of the 17th International Congress of Phonetic Sciences, Hong Kong, August*, str. 1694–1697.
59. Rose, P. (2002): *Forensic Speaker Identification*. London: Taylor & Francis.
60. San Segundo, E. i Gomez-Vilda, P. (2014): Evaluating the forensic importance of glottal source features through the voice analysis of twins and non-twin siblings. *Language and Law*, 1(2), 22-41.
61. Schiller, N.O. i Koster, O. (1996): Evaluation of a foreign speaker in forensic phonetics: a report. *Forensic Linguistics*, 3(1), 176-185.
62. Shipp, T., Doherty, E.T. (1987): Some fundamental considerations regarding voice identification. *Journal of the Acoustical Society of America*, 82, 687-688.
63. Tanner, D.C. (2007): *Medical-legal and Forensic Aspects of Communication Disorders, Voice Prints, and Speaker Profiling*. Tucson, AZ: Lawyers & Judges Publishing Company.
64. Tosi, O., Oyer, H., Lashbrook, W., Pedrey, C., Nicol, J., Nash, E. (1972): Experiment on Voice Identification. *The Journal of the Acoustical Society of America*, 51(6), 2030-2043.

65. Traunmuller, H., Eriksson, A. (1994): The frequency range of the voice fundamental in the speech of male and female adults. Unpublished manuscript, Institutionen för lingvistik, Stockholms universitet.
66. Traunmuller, H., Eriksson, A. (2000): Acoustic effects of variation in vocal effort by men, women, and children. The Journal of the Acoustical Society of America, 107, 3438–3451.
67. Zatorre, R.J., Belin, P. (2001): Spectral and temporal processing in human auditory cortex. Cerebral cortex, 11(10), 946-953.
68. Young, M.A. i Campbell, R.A. (1967): Effects of Context on Talker Identification. The Journal of the Acoustical Society of America, 42(6), 1250.

### **Elektronički izvori**

- Boersma, P. and Weenink, D. (2012). Praat: doing phonetics by computer (version 5.3.79). Pristupljeno 5.srpnja.2017. <http://www.fon.hum.uva.nl/praat/>
- BioMet®Soft, (2010). Biomet®soft. universidad politécnica de madrid. Pristupljeno 27.6.2017. <http://www.glottex.com/en/component/content/article?id=2:cursillo-de-formacion-en-evaluacion-de-calidad-de-voz-para-logopedia-con-biomet-phon>
- van Lieshout, P. (2003): PRAAT: Short tutorial. University of Toronto, Graduate Department of Speech-Language Pathology, Faculty of Medicine, Oral Dynamics Lab (ODL). Pristupljeno 13. lipnja, 2017. [https://web.stanford.edu/dept/linguistics/corpora/material/PRAAT\\_workshop\\_manual\\_v421.pdf](https://web.stanford.edu/dept/linguistics/corpora/material/PRAAT_workshop_manual_v421.pdf)
- The International Association for Forensic Phonetics and Acoustics (IAFPA), Helsinki, 2004; Pristupljeno 9.ožujka, 2017. <http://www.iafpa.net/code.htm>  
[https://www.bka.de/EN/OurTasks/SupportOfInvestigationAndPrevention/ForensicScience/PhysicalEvidence/Extortion/SpeakerIdentification/speakeridentification\\_node.html](https://www.bka.de/EN/OurTasks/SupportOfInvestigationAndPrevention/ForensicScience/PhysicalEvidence/Extortion/SpeakerIdentification/speakeridentification_node.html) Pristupljeno 13. ožujka, 2017.
- <http://www.erf.unizg.hr/hr/studiji/preddiplomski-logopedija> Pristupljeno 9. lipnja, 2017.
- <http://www.fon.hum.uva.nl/praat/manual/Scripting.html> Pristupljeno 14., srpnja, 2017.
- <http://www.taracentar.hr/o-nama/forenzicna-fonetika-i-akustika/> Pristupljeno 14. srpnja, 2017.



<https://tkojetko.irb.hr/znanstvenikDetalji.php?sifznan=4428&podaci=projekti&sifprojekat=4882> Pristupljeno 7. rujna, 2017.

[http://www.ffzg.unizg.hr/fonet/?page\\_id=122](http://www.ffzg.unizg.hr/fonet/?page_id=122) Pristupljeno 7. rujna, 2017.

<https://bib.irb.hr/lista-radova?autor=153544> Pristupljeno 7. rujna 2017.

## POPIS SLIKA I TABLICA

Slika 1: Spektrogram (sonogram) (Preuzeto iz Heðever, 2012). .....	6
Slika 2: Grafički prikaz fundamentalne frekvencije (F0) (crveni trag) i amplitude (plavi trag) između dva ciklusa, prikazano programom MDVP (Preuzeto iz KayPENTAX, 2008).....	14
Slika 3: Radijalni dijagram MDVP programa s prikazom parametara (Preuzeto iz KayPENTAX, 2008). ..	15
Slika 4: MDVP izvješće dobivenih glasovnih parametara (Preuzeto iz KayPENTAX, 2008). .....	16
Slika 5: Oscilogram (gornji dio prozora) i spektrogram (donji dio prozora) dobiveni programom PRAAT (Preuzeto iz Lieshout, 2003). .....	19
Slika 6: Grafički prikaz jittera (frekvencijskih perturbacija) i shimmera (amplitudnih perturbacija) (Preuzeto iz Heðever, 2012). .....	25
Tablica 1: Auditivno uočljiva obilježja glasa i govora uz pripadajuće kategorije.....	12
Tablica 2: Popis parametara unutar MDVP programa (Preuzeto iz KayPENTAX, 2008).....	16